NHGRI at the ASHG 61st Annual Meeting
October 11-15, 2011
Montreal, Canada

DAY 2
WEDNESDAY
OCTOBER 12, 2011

Plenary Talk

8-10 a.m.
Plenary Session on Epigenetics
Room 210, Level 2, Convention Center

9:25-9:40 a.m.
Program Number 1

Identification of a mosaic activating mutation as the molecular basis of Proteus syndrome using massively parallel sequencing of affected tissues

L. G. Biesecker, *et al*
Includes from NHGRI: L. G. Biesecker, M. J. Lindhurst, J. C. Sapp, J. K. Teer, J. J. Johnston, K. Peters, J. Turner, P. L. Schwartzberg, J. C. Mullikin

Proteus syndrome manifests mosaic overgrowth in skin, connective tissue, brain, and other tissues. It was hypothesized to be caused by a mosaic mutation, lethal in the non-mosaic state. We tested this hypothesis using massively parallel sequencing of DNA from Proteus syndrome tissues, comparing affected to unaffected tissues. Seventeen samples from 11 patients were sequenced using targeted exome capture. These samples included four affected-normal sample pairs, two affected samples from a fifth patient, one pair of samples from discordant monozygotic twins, and five unaffected parents. We generated ~1.6 Gb of sequence data, with >87% coverage of exons with a high quality genotype. A novel variant was initially identified in a single affected sample using an affected-unaffected filter. Upon manual examination, we found this variant in three additional affecteds but not in the parents of affecteds or the unaffected identical twin, 1000 genomes (n=634), or ClinSeq (n=401). The sequence variant data were confirmed and extended by a custom restriction enzyme assay of >150 samples from 29 patients. Most affected samples were specimens removed from clinically abnormal areas at surgery. Twenty-seven of 29 patients with Proteus syndrome were found to have the identical mutation in this gene, but it was not present in >20 cell lines and tissues from persons with unrelated disorders. Tissues and cell lines from patients with Proteus syndrome harbored admixtures of mutant alleles that varied from 1% to ~50%. Patient-derived cell lines with the mutation showed evidence of activation of this gene product using anti-phosphoprotein antibodies and western blot analyses. We also show that a pair of single cell clones established from the same starting culture and differing only by their mutation status had differential activation of this protein. We conclude that a somatic mosaic activating mutation in this gene causes Proteus syndrome, validating the Happle mosaicism hypothesis. That this mutation is the same in all patients, mosaic, and not in a CpG dinucleotide explains the rarity of this disorder. These data show the power of massively parallel sequencing to identify causative genes in disorders that are not amenable to positional cloning. Further, these results provide a therapeutic target for this severe and progressive disorder. (The author confirms that the gene and mutation will be disclosed at the ICHG meeting, should the abstract be selected).

Platform Talks

4:15 p.m.-6:15 p.m.
Session 18 on Biochemical Genetics
Room 710B, Level 7, Convention Center

4:15 p.m.-04:30 p.m.
Program Number 67

***NBEAL2* is mutated in Gray Platelet Syndrome and is required for biogenesis of platelet alpha-granules.**
T. Vilboux, *et al*
Includes from NHGRI: C. F. Boerkoel, Y. Huang, D. Maynard, H. Dorward, K. Berger, J. C. Mullikin, M. Huizing, W. A. Gahl, M. Gunay- Aygun

Gray Platelet Syndrome (GPS) is a rare autosomal recessive disorder characterized by bleeding tendency, myelofibrosis, thrombocytopenia, and large platelets that lack α-granules. The causative gene has been sought for decades. We mapped the locus for GPS to a 9.4Mb interval on 3p21.1-22.1 that included 197 protein-coding genes. We sequenced these genes using a combination of next generation and Sanger sequencing in 15 independent GPS families. We identified 15 different mutations in *NBEAL2* (neurobeachin-like 2); 5 missense, 4 frameshift, 3 nonsense and 3 splice site mutations. The protein encoded by *NBEAL2* has no known function, yet; however, it belongs to a family of proteins that contains 3 domains (BEACH, ARM and WD40) that are crucial for protein-protein interactions, membrane dynamics and vesicle trafficking. Another protein from this family, LYST, a lysosomal trafficking regulator protein, is defective in Chédiak-Higashi syndrome, a disorder associated with platelet dense granule abnormalities in addition to giant secretory granules in leukocytes and other cell types. RNA analysis showed that at least 7 *NBEAL2* mRNA transcripts are expressed in hematopoietic cells, including megakaryocytes and platelets. Mass spectrometry of discontinuous sucrose gradient platelet fractions localized NBEAL2 protein to the platelet dense tubular system (endoplasmic reticulum). Microarray data in GPS fibroblasts showed overexpression of fibronectin, essential for proplatelet formation in cultured megakaryocytes and critical for megakaryocyte-matrix interactions. This could explain the myelofibrosis seen in GPS patients and therefore can be explored as a therapeutic target. Understanding NBEAL2 function will likely lead to the discovery of novel pathways of organelle formation and maturation.

You may contact the first author (during and after the meeting) at vilbouxt@mail.nih.gov

---

Session 17 on Molecular Basis I: Skin and Inflammatory Disorders
Room 710A, Level 7, Convention Center

4:30 p.m.-4:45 p.m
Program Number 60

Clinical transcriptome sequencing and analysis of a patient with spiny follicular hyperkeratoses of unknown etiology.
K. V. Fuentes Fajardo, *et al*
Includes from NHGRI: M. Huizing, M. Nehrebecky, F. Gill, C. F. Boerkoel, A. R. Cullinane, W. A. Gahl

Advances in Next Generation Sequencing technology (NGS) have made transcriptome sequencing a practical reality for clinical research applications. Exome sequencing as a tool for finding genetic causes of diseases has become prominent in the literature, and is chosen as the major NGS technology when following up cases in the NIH Undiagnosed Diseases Program (UDP) Protocol. To date, no diagnosis has been arrived at through leads from transcriptome research. The UDP's first attempt at transcriptome sequencing in a clinical setting involved a 50 year-old Caucasian woman with a unique phenotype that included widespread spiny follicular hyperkeratosis resulting in scarring alopecia, follicular plugging and skin abscesses. Her initial milder lesions were exacerbated by a combination of UV-A light treatment and oral retinoids. Thorough dermatologic and medical evaluations to rule out infectious, endocrinologic and paraneoplastic etiologies were unhelpful in delineating an etiology and in clearing the skin problem. Histology and candidate gene sequencing were not diagnostic. Here we present the analysis of transcriptome sequence variations found in RNA from keratinocytes and fibroblasts from involved skin contrasting and comparing them to those found in RNA from fibroblasts and keratinocytes from unaffected skin, blood, and the reference human genome sequence. We identified 18 mutations unique to the keratinocytes and fibroblasts from affected areas, including two genes that indicate a "second-hit," causing additional damage to a gene already harboring a missense mutation. We also observed that deleterious germline mutations were enriched for genes in cell adhesion, cell transport and immune response (p=0.00002, Bonferroni). These data demonstrate the ability to use RNA-Seq to find genes that carry deleterious mutations and demonstrate dysregulation of expression. These findings can then guide follow-up targets for diagnostic and treatment decisions.

You may contact the first author (during and after the meeting) at Karin.FuentesFajardo@nih.gov

Session 18 on Biochemical Genetics
Room 710B, Level 7, Convention Center

5:30 p.m.-5:45 p.m.
Program Number 72

Whole exome sequencing identifies a novel mitochondrial enzyme as the gene responsible for Combined Malonic and Methylmalonic Aciduria.

J. L. Sloan, *et al*
Included from NHGRI: J. J. Johnston, I. Manoli, R. J. Chandler, C. Krause, N. Carrillo-Carrasco, S. D. Chandrasekaran, J. R. Sysol, K. O'Brien, N. S. Hauser, J. C. Sapp, H. M. Dorward, M. Huizing, N. I. H. Intramural Sequencing Center, L. G. Biesecker, C. P. Venditti

Combined methylmalonic and malonic aciduria (CM a.m.MA) is characterized by increased urinary methylmalonic acid (MMA) and malonic acid (MA), with MMA>MA, and normal malonyl-CoA decarboxylase activity. CM a.m.MA was first reported in a child with failure to thrive, seizures and immunodeficiency and a dog with neurodegeneration but the molecular etiology was unknown. We performed exome sequencing of a patient with CM a.m.MA and her parents and filtered variants to identify 12 candidate genes. One of these genes had a putative mitochondrial leader. We sequenced this gene in 9 patients and found two mutations in this gene in 8 of 9 patients. DNA from the dog also had a homozygous mutation in this gene. Mutations included 9 missense, 1 in-frame deletion and 1 nonsense mutation. Eight missense mutations and the in-frame deletion were in functional motifs conserved among members of this enzyme family. One patient with 2 mutations was identified in an exome cohort of subjects not ascertained for metabolic disease and had the distinct biochemical features of CM a.m.MA. The age of diagnosis and symptoms in the nine subjects with CM a.m.MA were highly variable. MMA and MA were elevated in plasma and urine using a new GC-MS assay developed to measure MA. Fibroblasts from 4 subjects had a cellular metabolic defect; increased production of MMA (2.4- to 6-fold) compared to controls, after loading with 5 mM propionate. Viral overexpression of the candidate gene, but not GFP, corrected the metabolic defect. Immunostaining of fibroblasts overexpressing a C-terminal GFP fusion protein or the native enzyme showed a mitochondrial distribution and co-localized with a mitochondrial antibody. These data establish the causative gene for CM a.m.MA and describe the first disease association with a member of this enzyme family. Mutant alleles occur with a minor allele frequency (MAF) of 0.0058 in ~1,000 control individuals predicting a CM a.m.MA population incidence of ~1:30,000. This predicts that CM a.m.MA is one of the most common forms of MMAemia, and perhaps, one of the more common inborn errors of metabolism. The spectrum of symptoms and natural history of this disorder are highly variable and require further delineation. The identification of an affected using exome sequencing highlights an interesting and alternative diagnostic approach because CM a.m.MA is not identified through routine newborn screening by elevated propionylcarnitine.

You may contact the first author (during and after the meeting) at jsloan@mail.nih.gov

Poster Session

10 a.m.-7 p.m.
Exhibit Hall, Level 2, Convention Center

2-3 p.m.
Poster category: Ethical, Legal, Social and Policy Issues in Genetics

Program Number 1149W

National Institutes of Health Controlled Data Access: Experience of the Genetic Association Information Network Data Access Committee.
C. Din-Lovinescu, *et al*
Includes from NHGRI: E. Bookman, L. Rodriguez, E. Ramos

Genome Wide Association Studies (GWAS), a strategy for identifying genetic variants associated with health and disease, produce vast amounts of data suitable for addressing many research questions. To maximize the scientific uses of GWAS data through broad data sharing, the National Institutes of Health (NIH) developed a controlled-access data repository and data access system designed to provide GWAS data to qualified investigators for research uses consistent with the informed consent of study participants. The Genetic Association Information Network (GAIN), one of the first NIH GWAS programs, has been at the forefront of rapid and broad data sharing since its inception in 2006 and helped shape the development and design of the database of Genotypes and Phenotypes (dbGaP) and its data access system. GAIN includes GWAS of ADHD, diabetic nephropathy, psoriasis, major depression, schizophrenia, and bipolar disorder. Data Access Requests (DARs) for GAIN studies are reviewed by the GAIN Data Access Committee (DAC). The DAC grants access to controlled-access data after ensuring that investigators have adequate credentials and meet the requirements of the data access request process, including proposing research that is consistent with the data use limitations of the requested data. As of April 30, 2011, 808 DARs have been submitted, 654 (81%) of which were approved with an average submission to decision time of 12.9 days. The schizophrenia and bipolar disorder datasets are requested most frequently. 57% of submitted DARs include schizophrenia; 46% include bipolar disorder. The most common proposed research use is to study GAIN-specific phenotypes (53% of submitted requests), followed by methods development (18%) and adding controls to other studies (17%). 316 investigators have been approved to access GAIN data with 165 resulting publications. Eight investigators approved to access GAIN data (2.5%) were involved in a data management incident where the terms of the NIH Data Use Certification agreement were not followed. The majority of these incidents occurred in the early years of dbGaP and all were resolved. As the number of studies deposited in dbGaP continues to grow, a summary of how GAIN data are being accessed and used along with the experience of the GAIN DAC can serve as an example to inform the NIH and the research community about the impact of making GWAS and other large-scale genomics datasets broadly available.

You may contact the first author (during and after the meeting) at corina.din-lovinescu@nih.gov

---

2-3 p.m.
Poster category: Molecular Basis of Mendelian Disorders

Program Number 1135W

Mitochondrial mistargeting causes autosomal dominant renal Fanconi syndrome.
E. Klootwijk, *et al*
Includes from NHGRI: A. Helip-Wooley, W. A. Gahl

Renal Fanconi syndromes are characterized by a generalized renal proximal tubular dysfunction leading to aminoaciduria, glucosuria, phosphaturia, small molecular weight proteinuria, and metabolic acidosis. The clinical consequences in childhood include rickets and growth problems. We investigated the genetic basis and the underlying molecular pathology in an extended family with autosomal dominantly inherited renal Fanconi syndrome without kidney failure. Whole genome multipoint parametric linkage analysis was performed, revealing a significant LOD score (> 3) for a single locus. All genes in the linked area were sequenced resulting in the identification of a heterozygous mutation in a gene, which we call Fanconin; the mutation leads to de-novo formation of a mitochondrial targeting motif. This mutation was not found in ethnically matched controls. To assess its functional impact, a stable permanently transfected inducible renal proximal tubular cell model was generated. Immunohistochemical analysis using these cells showed appropriate intracellular localization of the wild type human Fanconin, but mitochondrial mistargeting of the mutant human Fanconin. Immunohistochemical studies of mouse kidney showed that the Fanconin protein is expressed in renal proximal tubules, consistent with the proximal tubular dysfunction observed in our patients. Knockout mice for Fanconin did not exhibit a proximal tubular transport defect, as assessed by aminoacid analysis using GC-MS. This was consistent with the hypothesis that the renal Fanconi phenotype in our family is not caused by haploinsufficiency, but rather by a dominant negative effect of the Fanconin protein. Since Fanconin is mistargeted to mitochondria, we studied the impact of this mutation on mitochondrial function. Oximetric

analyses showed a complex coupled ATP synthase deficiency (oxidative phosphorylation) in cells expressing the mutant Fanconin. Transport studies documented abrogated transepithelial water transport. To our knowledge this is the first description of a genetic defect leading to intracellular mistargeting of a mutant protein resulting in mitochondrial pathology.

2-3 p.m.
Poster category: Metabolic Disorders

Program Number 1295W,

Identification of a novel cause of autosomal dominant, adult-onset distal myopathy.
M. C. Malicdan, *et al*
Includes from NHGRI: M. C. Malicdan, C. F. Boerkoel, Y. Huang, C. Groden, W. A. Gahl, C. Toro

Distal myopathies are genetically heterogenous diseases. Associated disease genes are known to encode proteins in the sarcomere (titin, myosin), plasma membrane (dysferlin, caveolin), or cytoskeleton (desmin, myotilin, αB-crystallin, ZASP, filamin C, and nebulin). Despite this progress, the majority of cases remain undefined and a systematic explanation for a disease mechanism is lacking. Only one cause of adult-onset distal myopathy affecting predominantly the posterior compartment has been genetically characterized. We define a second genetic cause in a two-generation family. Affected individuals have progressive weakness of distal muscles beginning late in the third decade of life without evidence of neuropathy, increased serum creatine kinase, respiratory involvement, or cardiac symptoms. MRI studies showed fatty replacement of the posterior compartments of the legs, mainly of the medial gastrocnemius. Muscle pathology showed variation of fiber size due to fiber atrophy and esoinophilic cytoplasmic inclusions that were devoid of desmin, plasma membrane proteins and oxidative activity. There were no rimmed vacuoles. Ultrastructural studies showed non-membrane bound fibrillary deposits 6-12 microns in diameter in the Z-bands. Whole exome and Sanger sequencing excluded mutations in known genes associated with distal myopathy and identified an intronic splice site mutation causing exon skipping and encoding a frameshift mutation in an unstudied enzyme. Based upon the dominant inheritance and comparable expression of the mutant and wild type mRNAs, the truncated enzyme is an antimorph or neomorph. We present spatiotemporal expression analyses and elucidate the cellular function of the enzyme and the potential disease mechanism. You may contact the first author (during and after the meeting) at malicdanm@mail.nih.gov

---

2 p.m.-3 p.m.
Poster category: Clinical Genetics and Dysmorphology

Program Number 987W

Case-parent trio genome-wide association study identifies several candidate loci for nonsyndromic sagittal craniosynostosis.
C. M. Justice, *et al*
Includes from NHGRI: Y. Kim, A. McMullen, H. Ling, A. F. Wilson, S. A. Boyadjiev

Craniosynostosis is a common malformation in which one or more of the cranial sutures in an infant skull (metopic suture, coronal sutures, sagittal suture and lambdoid sutures) fuse prematurely. Sagittal craniosynostosis is the most common type of craniosynostosis, accounting for 40 to 58% of all cases, with an estimated prevalence of 1.9-2.3 per 10,000 live births. The International Craniosynostosis Consortium (genetics.ucdmc.ucdavis.edu/icc.cfm) has been established with the aim of identifying the genetic causes of nonsyndromic craniosynostosis and more than 720 families with at least one affected individual have been recruited. In an attempt to identify genetic variants associated with nonsyndromic sagittal craniosynostosis, we performed a genome-wide association analysis of 201 case-parent trios and 13 nuclear families with two affected siblings. A total of 662 individuals in 214 families (70% Caucasian families, 30% mixed ethnicity families) were genotyped on the Illumina 1M Human Omni1-Quad array. Quality control measures reduced the number of SNPs available for analysis from 1,140,419 to 914,402 markers. Association between sagittal craniosynostosis and each SNP was measured using the transmission disequilibrium test (TDT) as implemented by PLINK v1.0.7. The strongest association was with rs1884302 ($p = 3.79 \times 10^{-14}$) in the flanking 3' UTR of BMP2 on chromosome 20. BMP2 belongs to the transforming growth factor-beta (TGFB) gene family and is involved in bone and cartilage formation. Genome-wide significant ($p \leq 5 \times 10^{-8}$) associations were also detected for SNPs intronic to BBS9 (rs1420154, $p=3.51 \times 10^{-13}$) on chromosome 7, which is thought to be involved in parathyroid hormone action in bones. Additionally, we have strong but not genome-wide significant signals on DLG1 (rs12152266, $p=1.44 \times 10^{-7}$) on chromosome 3, RPS12 (rs9493468, $p=2.8 \times 10^{-7}$) on chromosome 6 and LOC643631 (rs1948330, $p=4.78 \times 10^{-7}$) on chromosome 5. Loss of heterozygosity mapping identified several chromosomal regions which are in process of further analysis.
You may contact the first author (during and after the meeting) at cmj@mail.nih.gov

---

3 p.m.-4 p.m.
Poster category: Clinical Genetics and Dysmorphology

Program Number 1002W

Genetic Medicine: The NIH Undiagnosed Diseases Program Model.

C. F. Boerkoel II
Includes from NHGRI: D. Adams, T. Markello, C. Toro, C. J. Tifft, W. A. Gahl

In classical evolutionary theory, human disease arises by maladaptation of humans to their ecological niche. By definition common diseases arise in response to abrupt changes in the ecological niche exposing genetic variants underlying susceptibility and resistance to the ecological change. Modifying the environment therefore likely best treats common diseases. In contrast, rare diseases, which affect ~8% of the population, frequently arise from strong genetic and epigenetic mutations causing maladaptation within a stable niche. Therapy for rare diseases therefore targets modification of the dysfunctional physiological pathway. Within this context, a precise diagnosis is the first step to understanding illness and defining appropriate therapies. Focusing on ill individuals without diagnoses, the NIH Undiagnosed Diseases Program (UDP) models a synthetic approach to diagnosis involving participatory care, transdisciplinary clinical evaluations, and integration of basic science tools into the clinical diagnostic paradigm. As a consequence of this "research is care approach, the UDP has been able to provide diagnoses to 15-20% of the patients seen. As a prelude for export of this model to the wider community, we have defined factors most influencing diagnostic ability, mechanisms allowing efficient integration of basic science tools into clinical care, and possible conceptual constructs for effective therapy in the future. In summary, we find that the infrastructure needed for personalized medicine is achievable despite several remaining conceptual hurdles. You may contact the first author (during and after the meeting) at boerkoelcf@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Statistical Genetics and Genetic Epidemiology

Program Number 630W

Phenotype-Genotype Integrator (PheGenI): Synthesizing Genome-Wide Association Study Data with Existing Genomic Resources.

E. M. Ramos
Includes from NHGRI: H. A. Junkins, L. A. Hindorff

Objective: Rapidly accumulating data from genome-wide association studies (GWAS) and other large-scale studies will be most useful when synthesized with existing databases. To address this, we developed the Phenotype-Genotype Integrator (PheGenI), a user-friendly web interface that integrates various National Center for Biotechnology Information (NCBI) genomic databases with association data from the NHGRI GWAS Catalog and supports downloads of search results. **Background:** The GWAS design has identified over 4,400 genetic variants associated with over 200 human traits and diseases. Rarely are the functional consequences of these variants understood, and replication, functional and follow up studies are crucial next steps. Integration of GWAS data with existing complementary databases can facilitate prioritization of variants to follow up, study design considerations, and generation of biological hypotheses. **Methods:** A number of existing genomic resources are housed at the NCBI, including dbSNP, Gene, and the Genotype-Tissue Expression (GTEx) program. GWAS association data are available through two other NIH resources, the Database of Genotypes and Phenotypes (dbGaP) and the NHGRI GWAS catalog. The availability of these centralized resources made possible the development of PheGenI (http://www.ncbi.nlm.nih.gov/gap/PheGenI), an online portal for scientists and clinicians who use or produce GWAS data, to browse, search, integrate and download results. **Results:** 7,614 association records have been integrated with approximately 38,000 records from dbSNP, 46,000 records from Gene, and 61,000 expression QTL records from GTEx. After weighting for associations that belong to multiple trait categories, 78% of these association records are distributed among a few categories: Chemicals and Drugs, Digestive System Diseases, Eye Diseases, Immune System Diseases, Mental Disorders, Neoplasms, Nervous System Diseases, Skin and Connective Tissue Diseases, and Other. PheGenI features include the ability to search by phenotype, gene, SNP, or chromosomal range. In addition, users can filter search results by association p-value or SNP functional class; display results locally on DNA sequence tracks or globally across the genome; and download data tables. **Conclusions:** PheGenI will facilitate follow up of results from GWAS, enabling scientists to investigate functionality of identified SNPs and interrogate relationships between SNPs and human disease.

You may contact the first author (during and after the meeting) at ramoser@mail.nih.gov

Program Number 646W

High-resolution array-CGH analysis of germline DNA in a melanoma-prone family linked to chromosome 9p21.
R. Yang, *et al*
Includes from NHGRI: J. Struewing

Germline copy number variations (CNVs) have recently been recognized as a significant source of genetic variation and have been associated with disease susceptibility. Chromosome 9p21 has been implicated in the pathogenesis of cutaneous malignant melanoma (CMM) because it contains CDKN2A, the major known high-risk CMM susceptibility gene. However, a subset of CMM families linked to 9p21 has no CDKN2A mutations. The goal of this study was to determine whether CNVs, particularly 9p21 CNVs, were related to CMM risk in a 9p21-linked melanoma-prone family without a CDKN2A mutation. The family has 13 CMM members and 6/7 genotyped CMM cases shared a common 9p21 haplotype. An extensive mutation analysis did not detect any mutations in coding or non-coding regions of CDKN2A. We therefore conducted a genome-wide search for CNVs using the Nimblegen 385K whole-genome array-CGH. We analyzed genomic DNA from 4 CMM cases and 1 spouse. We used the Nexus Copy Number ™ built-in Rank Segmentation algorithm to identify significant CNVs (significant threshold=0.000001; minimal number of probes per segment=5; log2 ratio>0.2 for gains and -0.3 for losses). No CNVs were consistently shared by multiple CMM cases. However, there were some small CNVs in the 9p21 linked region that occurred in some but not all CMM cases. Given the importance of 9p21 in melanoma susceptibility and the observed linkage to this region in this family, we further evaluated CNVs in this region using two independent approaches: 1) a custom-made array-CGH design to focus on the linked region (average probe spacing, 13 bp) [analyzing 3 cases]; and 2) quantitative PCR (qPCR) targeting 4 genes (SH3GL2, CDKN2A, KIAA1797, and SLC24A2) in this region [analyzing all 26 individuals with available DNA]. No CNVs that were shared by multiple cases were identified suggesting that CNVs in the 9p21 region do not cause melanoma predisposition in this family. We are currently using next-generation sequencing technology to examine all exons in the 9p21 linkage region as well as exomic and complete genomic sequencing to identify disease-related variant(s) in the entire exome/genome in this family.

You may contact the first author (during and after the meeting) at royang@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Genomics

Program Number 804W

Optimized Filtering of High Throughput Sequence Variants in Clinical Cases from the National Institutes of Health Undiagnosed Diseases Program.
D. R. Adams, *et al*
Includes from NHGRI: M. Sincan, K. Fuentes Fajardo, C. Toro, C. F. Boerkoel, C. J. Tifft, W. A. Gahl, T. C. Markello, NIH Intramural Sequencing Center
High Throughput Sequencing (HTS) of exomes and genomes is rapidly becoming incorporated into the standard armamentarium of basic scientists and is making inroads into clinical molecular genetics. Several publications have discussed the data acquisition phase of HTS including alignments, base calling and annotation. The challenge of sorting through the resulting large list of candidate DNA variations is also well known, but has been evaluated to a lesser extent in the literature. We have explored the various means of filtering and sorting HTS variant lists using data from the NIH Undiagnosed Diseases Program (UDP). We present an analysis of the methodology and stringency considerations at each data filtering step, and discuss the integration of HTS data with genetic and clinical information from other sources. For this study, 30 probands plus additional family members were subjected to HTS of their exomes or genomes. In each case, extensive clinical evaluation at the NIH Clinical Center was used to attempt diagnosis and perform phenotyping. Analysis of the HTS variants generated for each individual or family was performed using both published and novel techniques, including the incorporation of data from high-density SNP arrays. The presence of four successful, verified molecular diagnoses provided a positive control for the efficacy of filtering methods and strategies. Filtering procedures at each step were tuned to match the clinical information available for the family or individual. Marked improvements in HTS variant filtering corresponded with improvements in filtering methodology and other resources such as successive dbSNP releases. Careful design of each filtering step allowed for optimization of the exclusion of false-positive variants. At the same time, flexible parameterization of each step allowed for systematic stringency adjustments. Laboratory confirmation of a DNA sequence variants of unknown significance can be time consuming and expensive. Therefore, careful filtering of HTS variant lists is essential for excluding candidates that can be either ruled out or de-emphasized prior to experimental validation.

You may contact the first author (during and after the meeting) at dadams1@mail.nih.gov

DAY 3
THURSDAY
OCTOBER 13, 2011

Invited Talk

8 a.m.-10 a.m.
Session 24 on Emerging Ethical Issues in Large-Scale International Genomics Research Collaborations
Room 511, Level 5, Convention Center

8:25 a.m.-8:45 a.m.
Introduction. J. E. McEwen NHGRI/NIH

Ethical issues in international genomics research collaborations: Perspectives from Sub-Saharan Africa.
C. Rotimi NHGRI/NIH

Large-scale genomics research projects—especially those that involve plans for the broad release of samples and data—raise many ethical issues. These issues are heightened in the context of international collaborations (for example, the International HapMap Project, the 1000 Genomes Project, and The Cancer Genome Atlas, and the International Human Microbiome Consortium), where different sets of cultural values and norms, and different legal and regulatory requirements, may have to be reconciled. For example, approaches to recruitment, informed consent, and individual autonomy and privacy vary considerably in different parts of the world. Attitudes about the role (if any) for community consultation or engagement and the appropriateness of particular research governance mechanisms can also differ. In addition, people in different parts of the world (and often even in the same part of the world) may have very different views about whether, or how, research results or incidental findings should be returned to participants in genomics research studies (where the samples and data have not been anonymized). The scope of the right to withdraw samples or data from repositories or databases is also an area of potential difference—especially where full realization of a project's end goals depends on having the samples and data available through a central repository to multiple researchers over an extended period of time. This session will explore these and other ethical issues that arise in large-scale international genomics research collaborations. Perspectives will be presented from several continents and countries.

Platform Talk

10:15 a.m.-12:15 p.m
Session 37 on Therapy for Genetic Disorders
Room 710A, Level 7, Convention Center

10:30 a.m.-10:45 a.m.
Program Number 140

A liver-specific transgenic mouse model identifies new disease-associated biomarkers and establishes antioxidants as an ameliorative treatment for the renal disease of methylmalonic acidemia (MMA).
E. Manoli, *et al*
Includes from NHGRI: J. R. Sysol, J. L. Sloan, R. J. Chandler, C. P. Venditti

We generated mice that express methylmalonyl-CoA mutase (*Mut*) cDNA under the control of a liver-specific promoter on a knockout background (*Mut*-/-;TgINS-Alb-*Mut*) to model extrahepatic manifestations, study pathophysiology and examine therapeutic interventions. Low-level hepatic *Mut* expression conferred complete rescue from the neonatal lethality displayed by *Mut*-/- mice and allowed disease-associated renal pathology to be induced with a high-protein diet. Ingestion of a high-protein chow for 2 months resulted in elevated plasma methylmalonic acid levels (µM) in the *Mut*-/-;TgINS-Alb-*Mut* mice (1500 620) compared to similarly-treated *Mut*+/- littermates (7.4 ±0.6), growth failure and increased mortality. *Mut*-/-;TgINS-Alb-*Mut* mice developed tubulointerstitial nephritis associated with a decreased glomerular filtration rate (GFR) [37.6 3.9% of *Mut*+/- GFR, $p<0.0001$] and elevated creatinine levels. Mitochondria of proximal tubular epithelial cells were enlarged and had shortened cristae; kidney immunohistochemistry showed increased succinate dehydrogenase and decreased cytochrome c oxidase staining. Expression analysis using whole kidney RNA from the protein-challenged *Mut*-/-;TgINS-Alb-*Mut* mice compared to age, sex and diet matched littermates revealed differentially expressed mRNAs from several pathways including immune response, lipid metabolism, ketone biosynthesis and cell survival. One significantly up-regulated gene encoding a secreted glycoprotein was also increased in the plasma and urine of *Mut*-/-;TgINS-Alb-*Mut* mice and the concentration correlated inversely with GFR ($r=-0.45$; $p<0.01$). This protein was further studied in the plasma of 46 *mut* MMA patients (NCT00078078); levels were strongly associated with renal indices (patient plasma creatinine and cystatin-C values; $p<0.01$ for both). The inclusion of ubiquinol and vitamin E in the high protein diet for 2 months ameliorated the loss of GFR in the *Mut*-/-;TgINS-Alb-*Mut* mice (37.6 ±3.9% [pre-] vs. 60 ±4.8% of *Mut*+/- GFR [post-treatment], $p<0.01$) and

normalized urinary and plasma disease-associated biomarkers despite persisting metabolite elevations induced by the dietary challenge. This novel mouse model has allowed identification of new biomarkers using genomic approaches and provided evidence for a therapeutic effect of antioxidants on the renal disease of MMA. These results should have broad extension to other metabolic disorders manifesting mitochondrial dysfunction.

You may contact the first author (during and after the meeting) at manolii@mail.nih.gov

---

Poster Session

10 a.m.-7 p.m.
Exhibit Hall, Level 2, Convention Center

2 p.m.-3 p.m.
Poster Category: Complex Traits: Theory and Methods

Program Number 391T

A mutation in BMP3 Contributes to Canine Brachycephaly.
E. A. Ostrander
A. Byers, B. Carrington, D. Faden, R. Sood, J. W. Fondon, J. J. Schoenebeck

Cephalic disorders occur with a frequency greater than 1 in 2,500 live births, yet the genetic underpinnings of most cephalic disorders remain unknown. For many purebred dogs, skull shapes reminiscent of cephalic disorders such as brachycephaly, dolichocephaly, and hydrocephalus are breed-defining features. Geometric morphometric analysis of 348 purebred dog skulls from 93 breeds was used to quantify skull shape. After correcting for size, we used principal components analysis to characterize remaining shape variance. Fifty-four percent of variance was explained by the first principal component (PC1), whose factor loadings included changes in rostrum length, cranial vault depth, and width of the zygomatic arches. Thus, PC1 is a quantitative measure of morphological changes pertinent to brachycephaly (rostrocaudal skull shortening). A genome-wide association scan of PC1 revealed numerous, highly significant QTLs. Detailed analysis of the association on chr32: 8-8.5 Mb (best SNP, *P* <3x10-44) revealed a depression in heterozygosity (Ho) among brachycephalic dogs. Loss of Ho is a hallmark of selective sweeps and can be indicative of a locus experiencing strong, selective pressure over generations of planned breedings. Fine mapping on chr32 revealed a missense F->L mutation in the predicted signaling portion of canine *BMP3*. Among the TGF β superfamily, the amino acid mutated in *Bmp3* is invariably aromatic (Phe or Tyr), leading us to hypothesize that such a mutation could lead to loss of function. To test our hypothesis, we turned to zebrafish as an experimental surrogate for testing *Bmp3* function. Our data show that during development, zebrafish *bmp3* is expressed in neurocranial tissues. Furthermore, morpholino knockdown of *bmp3* led to loss of cartilaginous head structures. Overexpression of wild type and mutant canine Bmp3 RNA demonstrates that only the first is biologically active in fish. Together, these experiments reveal a conserved role for *Bmp3* during vertebrate cranioskeletal development and suggest that the canine F->L mutation renders this molecule inactive. Identification of *Bmp3* as a contributor to canine brachycephaly demonstrate the power of studying dog genetics to uncover factors and pathways likely to be relevant to human cephalic disorders, thus enabling genetic diagnostics and new treatment paradigms.

You may contact the first author (during and after the meeting) at eostrand@mail.nih.gov

---

2 p.m.-3 p.m.
Poster Category: Statistical Genetics and Genetic Epidemiology

Program Number 615T

Replication of GWAS signals and association of novel functional variants for serum uric acid and total serum bilirubin levels in an Irish population.
Y. Kim, *et al*
Included from NHGRI: C. D. Cropp, L. C. Brody, J. E. Bailey-Wilson, A. F. Wilson

Background: The measurement of blood and serum metabolites can provide insight into the metabolic processes underlying the variation of traits related to many diseases. The catabolic products of purine and hemoglobin metabolism can directly cause disease and serve as pathologic markers of several disease states. In order to better understand the genetic determinants of the product of these pathways, we measured serum uric acid (URIC, mmol/L) and total serum bilirubin (TBIL, umol/L) in a healthy, young population. Blood and dietary information were collected and metabolites were measured in 2490 healthy students at Trinity College Dublin in Ireland. **Methods:** URIC was adjusted for age, sex, and BMI, and log-transformed TBIL was adjusted for age and sex. High density genotyping was performed with the Illumina 1M HumanOmni1-Quad chip. After quality

control assessment 2232 unrelated individuals and 757,533 SNPs were retained for association testing. Tests of association were performed with simple univariate linear regression, regressing each marker on each trait, assuming an additive genetic model as implemented in PLINK v1.0.7. Locus-specific heritability (h2) for the effect size of each SNP on phenotypic variation was calculated using R v2.12.1. **Results:** For URIC, we replicated previously reported SNPs in *SLC2A9* (rs6449213, h2 = 10%), *WDR1* (rs717615, h2 = 3%), and *ABCG2* (rs2199936, h2 = 2%) at genome-wide significance levels (p-value ≤ 5e-08). Additional SNPs were significant in these 3 reported candidate genes and *ZNF518B*. The most significant results were from rs13111638 (p-value = 4e-24, h2 = 9.5%) in an intronic region and in two coding SNPs in *SLC2A9* (rs10939650 and rs13113918 (p-value = 2e-22, 1e-21; h2= 4.4%, 5.6% respectively). For TBIL, strong *UGT1A* gene family signals were found including replication of rs887829 (p-value = 4e-156, h2=28%). Additionally, two non-synonymous SNPs (rs6431631 and rs1500480 (p-value = 3e-09, 2e-08 respectively) were found to be associated in *LOC339766*. However, there were no genome-wide significant signals in *SLCO1B1* a known candidate gene for TBIL (minimum p-value = 7e-07). **Conclusion:** Previously published common variants associated with URIC and TBIL levels with high h2 were replicated and additional new coding genetic variants were found in this healthy, young Irish population. These findings provide additional evidence for future analyses to identify the mechanism of genetic contributions to these phenotypes.

You may contact the first author (during and after the meeting) at kimyoo@mail.nih.gov

---

2 p.m.-3 p.m.
Poster Category: Clinical Genetics and Dysmorphology

Program Number 987T

Characteristics of kidney and liver disease in 38 patients with Joubert syndrome and related disorders (JSRD).
J. De Dios, *et al*
Includes from NHGRI: T. Vilboux1, J. Bryant, M. Huizing, T. Heller, W. A. Gahl, M. Gunay-Aygun

Joubert syndrome and related disorders (JSRD) are a genetically and phenotypically heterogeneous group of ciliopathies defined based on the unique combination of midbrain and hindbrain abnormalities resulting in the pathognomonic molar tooth sign" on axial brain imaging. JSRD patients can have mutations in a variety of genes that are also associated with other ciliopathies. Similar to other ciliopathies, a subset of patients with JSRD develop hepatorenal fibrocystic disease. Under the NIH clinical protocol, "Clinical Investigations into the Kidney and Liver Disease in Autosomal Recessive Polycystic Kidney Disease (ARPKD)/Congenital Hepatic Fibrosis and other Ciliopathies" (ClinicalTrials.gov: NCT00068224), we evaluated 38 JSRD patients. All patients had the typical "molar tooth sign". Age at NIH evaluation ranged from 0.9 to 36.2 years (8.6 + 7.3). Four patients from 4 families carried the diagnosis of Senior-Loken syndrome (SLS) and 12 from 10 families were classified as COACH syndrome based on presence of liver involvement. There was history of oligohydramnios in 4 and polyhydramnios in 2 patients. In 5 patients, prenatal ultrasound showed enlarged hyperechoic kidneys, indistinguishable from ARPKD. Six individuals including 3 SLS and 3 COACH syndrome patients received kidney transplantation, one after the NIH visit. Age at kidney transplantation ranged from 4 to 13 years (7.5 + 3). High resolution ultrasound (HR-USG) of the kidneys was normal in 20 of the 33 patients with native kidneys (ages 0.9 to 36.2 years, mean 7.0 + 8.1). The most common ultrasound finding was diffusely increased echogenicity with loss of corticomedullary distinction with or without discrete cysts; kidney size was enlarged in 2, small in 1 and normal in others. Two patients had unilateral multicystic dysplastic kidney. Glomerular filtration rate (GFR) was less than 80 ml/min/1.73 m2 in 12 patients, including 2 who had normal ultrasounds and GFRs of 52 and 65 at ages 9 and 3.5 years, respectively. While all COACH patients (11.6 + 8.8 years) had elevated liver enzymes, review of records showed normal liver enzymes in the first years of life. Nine patients including 7 with COACH had splenomegaly suggesting portal hypertension. No patients showed liver cysts. JSRD patients should be monitored for kidney and liver disease. Next generation DNA sequencing is underway. Genotype phenotype correlations might enable prediction of hepatorenal disease.

You may contact the first author (during and after the meeting) at john.dedios@nih.gov

---

3 p.m.-4 p.m.
Poster Category: Genetic Counseling and Clinical Testing

Program Number 1058T

Validation of open-source colorectal cancer risk assessment software compatible with the U.S. Surgeon General's My Family Health Portrait tool.
W. G. Feero, *et al*
Includes from NHGRI: F. M. Facio, A. Linn, D. Barton, L. G. Biesecker

Background -Family history (FH) is an acknowledged means for identifying the heritable component of an individual's cancer risk. The U.S. Surgeon General's My Family Health Portrait (MFHP) tool has been validated for the collection of family history information, but currently does not provide cancer risk assessment. Here we report on a newly developed open-source software tool designed to screen for heritable colorectal cancer (CRC) risk that is compatible with MFHP. Methods- A CRC risk assessment tool was developed using 2009 NCCN and USPSTF CRC risk assessment guidelines. The tool was designed to use patient-entered family history information to dichotomize individuals into not-elevated or elevated CRC risk categories, and provides basic age and risk-appropriate recommendations for CRC screening to all groups. The tool was validated on 150 pedigrees consecutively derived from the ClinSeqTM population. Risk assessments derived from MFHP were compared to two alternative "gold standards" for detecting risk applied to the same pedigrees: the CDC's Family HealthwareTM (FHT) CRC risk algorithm; and evaluation by 3 expert cancer genetic counselors (GC). Results- The MFHP, FHT, and GC evaluations identified 27, 14, and 16 elevated-risk probands, respectively, in the cohort. There was substantial agreement between GC CRC risk evaluations (average weighted kappa = 0.69 ± 0.04); two of three counselors had to agree that risk was elevated for a pedigree to be scored as such. Using different "gold standards for the identifying CRC risk, sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV) were: MFHP versus FHT, 100%, 90%, 52% and 100%; MFHP versus GC, 81%, 90%, 48% and 98%; FHT versus GC, 75%, 98%, 86% and 97%. Conclusions- The newly created software compatible with MFHP provides a tool for screening for heritable CRC risk. The sensitivities and specificities for MFHP derived using GC pedigree review as the "gold standard" for risk detection are not dissimilar from those of other types of screening tools used in primary care. The new tool has a higher sensitivity but lower specificity than FHT when compared to expert GC evaluation of CRC risk. The apparent "false positive" rate of the software is high; further research is needed to determine if this is a result of actual false positives or differences in risk assessment guidelines used by MFHP, FHT and cancer GCs.

You may contact the first author (during and after the meeting) at feerow@mail.nih.gov

Program Number 1060T

Clinical Genomics Data Infrastructure and ClinVar.
U. Geigenmüller,*et al*
Included from NHGRI: K. Fuentes Fajardo

Both cost and technical difficulty of sequencing large regions of the genome are dropping to a level that enables its routine use in patient care. Limiting the adoption of large-scale sequencing in molecular diagnostics is the difficulty of assessing the clinical significance of detected variants. Each testing laboratory typically performs its own interpretation based on local primary data and expertise in addition to publicly available information. This approach prevents laboratories, physicians, and patients from capitalizing on community expertise and experience and may lead to inconsistent interpretations between different laboratories. Data on variants detected in testing laboratories should be made readily available to the community. Furthermore, it would benefit patients and their physicians to have ready access to each laboratory's interpretation for any given variant. Convergence among interpretations from different laboratories would increase confidence in the consensus interpretation, while divergence would indicate uncertainty and offer opportunity to sharpen the scoring algorithms. A group of major commercial and academic laboratories and bioinformatics groups have come together to support NCBI's ClinVar database as a 'pre-competitive' repository of DNA sequence variation detected in the course of clinical diagnostics. The group is establishing procedures and technologies for the sharing of variants, defining critical data parameters required to evaluate pathogenicity of these variants, and working with NCBI to ensure these data can be stored and freely exchanged through ClinVar. Also, a code of conduct is being developed to assure priority for matters relating to patient safety. By making all evidence that underlies variant interpretation freely available, participants in this initiative hope that one of the major barriers to widespread adoption of extensive DNA diagnostic sequencing will be reduced.

You may contact the first author (during and after the meeting) at ute.geigenmuller@channing.harvard.edu

---

3 p.m.-4 p.m.
Poster Category: Molecular Basis of Mendelian Disorders

Program Number 1130T

Whole Exome Sequencing in a single nuclear family finds novel compound heterozygous changes in KCTD7 causing late infancy-childhood neurodegeneration.
T. C. Markell,*et al*
Included from NHGRI: D. A. Adams, L. Wolfe, M. Sincan, K. Fuentes Fajardo, MP. Jones, U. Harper, S.

Chandrasekharappa, C. J. Tifft, C. Boerkoels, W. A. Gahl, Nisc Comparative Sequencing Program

Two of four siblings who were normal at birth developed late infantile ataxia with slowly progressive myoclonus and severe degenerative neurological dysfunction. Four grandparents, both parents and all four siblings underwent traditional linkage analysis with a maximum LOD score of 0.82 at 17% of the genome, which corresponded to the same regions located by Boolean logic SNP chip linkage. The parents and four siblings underwent whole exome sequencing with Agilent Sure select 38MB exon capture and sequencing on an Illumina GIIx platform with an average 150x coverage. For the proband 79% of the UCSC coding bases were covered with high quality, most probable genotype MPG score ≥ 10. A combined 112,936 variants were identified in the 6 genotypes. Using standard frequency, linkage, Mendelian consistency and deleterious prediction filters biallelic mutations in a single gene *KCTD7* were detected. This gene has been linked to the similar but non-identical phenotype myoclonic epilepsy in two previous families by homozygosity mapping. Our family extends the phenotype of recessive loss of function in *KCTD7*. It also demonstrates that SNP based linkage from simple Boolean logic can produce the same chromosome linkage map as traditional linkage programs run with microsatellite markers.

You may contact the first author (during and after the meeting) at markellot@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Statistical Genetics and Genetic Epidemiology

Program Number 616T

Exploration of pleiotropic effects of inflammation-related disease GWAS SNPs with C-reactive protein levels in the PAGE study.
J. D. Kocarnik,*et al*
Included from NHGRI: T. Manolio

Introduction: Inflammation is an important health measure related to many common complex diseases. C-reactive protein (CRP) is a circulating biomarker indicative of systemic inflammation. Genome-wide association studies (GWAS) have successfully identified several loci important to inflammation-related diseases including cardiovascular disease (CVD), diabetes and obesity. However, the associations between these variants and CRP have not been fully explored. We evaluated whether CRP levels are associated with previously identified variants from GWAS of inflammation-related diseases, and whether such associations differ by race. Methods: We included 9088 participants of the Women's Health Initiative who had baseline high-sensitivity CRP measurements and were genotyped as part of the Population Architecture using Genomics and Epidemiology (PAGE) Consortium: 6309 White, 1564 Black, 714 Hispanic, 418 Asian/Pacific Islander, and 83 American Indian women. We evaluated 95 single nucleotide polymorphisms (SNPs) previously shown to be associated with inflammation-related diseases. We used linear regression stratified by race to evaluate the association between serum CRP level and each SNP, using an additive genetic model adjusted for age and global ancestry, using the first two principal components from ancestry informative markers. We corrected for multiple comparisons using Bonferroni adjustment ($p<5.3x10-4$). Results: Three SNPs had statistically significant positive associations with CRP: rs429358 T allele (APOE) in Whites ($p=1.2x10-20$), Blacks ($p=2.6x10-7$), and Hispanics ($p=3.3$ x10-7); rs4420638 A allele (APOC1) in Whites ($p=7.7x10-14$) and Hispanics ($p=8.7$ x10-5); and rs1260326 T allele (GCKR) in Whites ($p=6.2$ x10-6). No SNPs were significantly associated with CRP in Asian/Pacific Islanders or American Indians, though sample sizes were small. Conclusions: This preliminary analysis indicates that several SNPs previously associated with inflammation-related disease are also associated with circulating CRP levels, which appear to differ by race/ethnicity. Future efforts will include incorporating data from the other three studies of PAGE, to further detect and replicate and SNP-inflammation associations in each ethnic group. PAGE includes the Multiethnic Cohort (MEC), Epidemiologic Architecture of Genes Linked to Environment (EAGLE), and the CALiCo Consortium (see pagestudy.org).

You may contact the first author (during and after the meeting) at kocarnik@u.washington.edu

Program Number 704T

The PhenX Toolkit: Facilitating Cross-study Analysis in Genomic Studies.
W. Huggins,*et al*
Included from NHGRI: H. Junkins, E. Ramos

Cross-study analyses of genome-wide association studies (GWAS) are needed to increase statistical power and replicate results. However, these analyses have been hindered by the lack of standard measures for risk factors, covariates, and outcomes that are shared by common, complex diseases. To address this need, the PhenX (consensus measures for Phenotypes and

eXposures) Toolkit (https://www.phenxtoolkit.org/) offers high-quality well-established measures of phenotypes and exposures for use by the scientific community. The Toolkit contains 295 measures drawn from 21 research domains (fields of research). The measures were selected by Working Groups of domain experts using a consensus process that included input from the scientific community. Because measures relevant to complex diseases (e.g., Cardiovascular and Diabetes) can be found in multiple domains, measures are also organized into Collections, which are groups of measures related to a specific topic or target population (e.g., Diet and Nutrition and Body Size and Composition). For each PhenX Measure, the Toolkit provides a description of the measure, the rationale for including the measure in the Toolkit, protocol(s) for collecting the measure, and supporting documentation. Users can browse by measures, domains, or collections or search the Toolkit using the Smart Query Tool (SQT). Once users have selected some measures, they can download a customized Data Collection Worksheet (DCW) that specifies what information needs to be collected and a Data Dictionary (DD) that describes each variable included in their DCW. To help researchers share data, PhenX measures and variables are being mapped to multiple data standards, including the Unified Medical Language System (UMLS) and Cancer Data Standards Registry and Repository (caDSR). To help researchers find studies with comparable data, PhenX measures and variables are being mapped to studies in the database of Genotypes and Phenotypes (dbGaP). In collaboration with dbGaP staff at the National Center for Biotechnology Information (NCBI), 13 studies from the Gene Environment Association Studies (GENEVA) consortium and 3 studies from the electronic Medical Records and Genomics (eMERGE) network have been mapped to PhenX measures and variables. These mappings are displayed in dbGaP and highlight opportunities for cross-study analysis for researchers who adopt PhenX measures. *Supported by: NHGRI, Award No. U01 HG004597-01.*.

You may contact the first author (during and after the meeting) at whuggins@rti.org

---

3 p.m.-4 p.m.
Poster Category: Genomics

Program Number 802T

The genetics of idiopathic membranous nephropathy elucidated by GWAS.
H. Stanescu, *et al*
Included from NHGRI: M. Arcos-Burgos

Idiopathic membranous nephropathy (IMN) is a major cause of nephrotic syndrome and kidney failure, the etiology of which is not fully understood. We investigated the genetic basis of biopsy-proven cases of IMN in Caucasians. Three independent genome-wide association studies (GWAS) utilizing single nucleotide polymorphisms (SNPs) were performed in patients from Caucasian populations (75 French, 146 Dutch, 335 British cases of IMN). Cases were compared with ethnically matched controls; population stratification, quality controls, and statistics were carried out according to standard criteria. In a joint analysis from the 556 cases studied, we identified significant alleles at two genomic loci associated with IMN. Chromosome 2q24 contains the M-type phospholipase A2 receptor gene PLA2R1 (SNP rs4664308, p=8.6x10-29) shown previously to be the target of an autoimmune response. Chromosome 6p21 contains the human leukocyte antigen complex (HLA) class II gene HLA-DQA1 (SNP rs2187668, p=8.0x10-93). The association to HLA-DQA1 was significant in all three populations (p=1.8x10-9, 5.6x10-27 and 5.2x10-36 in French, Dutch and British, respectively). The odds ratio for IMN with homozygosity for both risk alleles is 79 (35-178 (95% confidence interval)). We sequenced both loci to ascertain the basis for our GWAS findings. An HLA-DQA1 allele on chromosome 6p21 is most significantly associated with IMN in Caucasians. This allele may facilitate an autoimmune response against targets, such as variants of PLA2R1. Our findings, including sequencing of both loci, suggest a basis for the understanding of IMN and illuminate how powerful GWAS can be for the study of a rare disease.

Program Number 846T

Whole exome sequencing of populations of African ancestry.
A. Adeyemo, *et al*
Included from NHGRI: A. Doumatey, H. Huang, D. Shriner, J. Zhou, C. Rotimi

The application of high throughput sequencing approaches to human exomes and whole genomes promises to provide new insights into population variation, disease mapping and evolutionary biology. African populations play a unique role by virtue of belonging to some of the oldest human lineages and exhibiting the greatest genetic diversity of any continental population. In the present study, we conducted whole exome sequencing of ten individuals of African ancestry, comprising Ghanaians (three Akan, two Gaa-Adangbe), Nigerians (two Yoruba and one Igbo) and African Americans (two individuals from the Washington

DC region). To our knowledge, this is the first time that individuals from three of the ethnic groups (Akan, Gaa-Adangbe, Igbo) are undergoing whole exome sequencing (WES). Sequence enrichment was done using Nimblegen SeqCap EZ Exome v2.0 which targets a total of ~30,000 coding genes (~300,000 exons, total size 36.5 Mb, 44.1 Mb target region). Sequencing was done on an Illumina HiSeq 2000. Reads were mapped and aligned, then aligned data files were converted, sorted, and indexed using Samtools and Picard. The sequence quality scores were recalibrated with the Genome Analysis Toolkit (GATK). After identification and removal of duplicate reads, variants were identified with GATK's UnifiedGenotyper and DINDEL tools. The single-nucleotide variants were then filtered for the removal of low-quality variant calls with GATK's VariantFiltration walker. Mean coverage per sample was ~88X. The number of sites in the covered regions that differed from the NCBI Build 36 (hg18) reference ranged from 25,911 to 35,394 per individual. The number of novel SNPs fell from ~5500 per person with reference to dbSNP129 (the last pre-1000 Genomes version) to ~1600 per person (with reference to dbSNP132). This represents a ~70% reduction in the number of novel variants, demonstrating the enormous contribution of the 1000 Genomes Project and other large-scale sequencing projects. The individuals represented by the newly sequenced groups (Akan, Gaa-Adangbe, Igbo) had variants that were absent from the 1000 Genomes YRI reference. This suggests that more samples from more African populations would prove useful in characterizing genetic variation and utilizing this variation to map disease genes. New and ongoing genomic initiatives in Africa (e.g., H3Africa and MalariaGen) will benefit from strategic integration of whole exome and whole genome sequencing technologies.

You may contact the first author (during and after the meeting) at adeyemoa@mail.nih.gov

Program Number 898T

GWAS in Neurofibromatosis Type 1 (NF1): progress update.

A. Pemov, *et al*
H. Sung, A. F. Wilson, D. R. Stewart

Background. NF1 is known for phenotypic variable expressivity and limited correlation between genotype and phenotype. However, the cause of such variability is unknown. We hypothesize that this variable expressivity is genetically determined, in part, by elements other than variants in coding regions at the *NF1* locus. We tested this hypothesis with a genome-wide association study (GWAS) of quantitative NF1 phenotypes. **Methods.** A single observer quantified severity in multiple NF1 sub-phenotypes, including head circumference (OFC), and the number of café-au-lait macules (CALM), cutaneous neuorofibromas (cNF), Lisch nodules (LN) and cherry hemangiomas (CH). Germline DNA was hybridized on the Illumina Human OmniQuad-1M SNP genotyping array for 117 Caucasians (82 families, average family size = 1.4). Traditional test of association, ignoring family structure was performed on 605,630 SNPs (MAF >= 0.10) on the 22 autosomal chromosomes in an attempt to identify non-NF1 variants associated with sub-phenotypes. Each sub-phenotype was adjusted for selected covariates such as age, gender and eye color. **Results.** Several SNPs were identified for each studied NF1 clinical sub-phenotypes at a significance level of 10-5 reflecting the modest sample size. Significant SNPs for CH and LN sub-phenotypes resided in gene-free regions of the genome, while the SNPs for CALM, cNF and OFC were found inside *RPS6KA2*, *KCNJ6* and *FOXP2* genes, respectively. *RPS6KA2* encodes p90 ribosomal S6 kinase, which, like *NF1*, is a part of MAPK and mTOR pathways. *KCNJ6* encodes one of the subunits of inward rectifier potassium ion channels; it has been shown that inactivation of *NF1* in Schwann cells leads to increased levels of c a.m.P and as consequence to increased outward K+ current and cell proliferation. *FOXP2* is a transcription factor that controls expression of ~300 human genes and is highly expressed in developing brain; this gene is required for proper development of speech and language regions of the brain during embryogenesis **Conclusions.** We expect that GWAS of quantitative traits in NF1 to be a useful first step in understanding NF1 phenotypic diversity. Future analyses will include other NF1-related traits. Principal component analysis will be performed on these phenotypes to identify independent phenotypic components for analysis.

You may contact the first author (during and after the meeting) at apemov@mail.nih.gov

DAY 4
FRIDAY
OCTOBER 14, 2011

Platform Talk

4:15-6:15 p.m.
Platform Session 56 on Complex Traits I: Approaches and Methods
Room 511, Level 5, Convention Center

6-6:15 p.m.
Program Number 210

Exome sequencing to identify genes harboring rare variants that determine lung function decline in COPD.
R. A. Mathias, *et al*
Includes from NHGRI: Y. Kim

Rationale: A recent genome-wide association study (GWAS) has revealed several loci determining rate of lung function decline in European American individuals with COPD from the NHLBI-supported Lung Health Study (LHS). Here, we use an exome-sequencing approach to identify additional rare coding variants not adequately tagged by the genomewide marker array. Methods: A GWAS was conducted using the Illumina 660W chip on 4,102 subjects with a minimum of 3 time-points where lung function was measured over a 5-year period. As part of the NHLBI-supported Exome Sequencing Program, a subset of these samples with extreme quartiles of lung function decline have been exome sequenced using Nimblegen capture arrays on the Illumina platform. Sequences were aligned using BWA, variants called using the Genome Analysis Toolkit, and annotated using Seattle-Seq. Standard sequencing quality control filters were applied. A case-control analysis was performed restricting several alternative burden tests to functional (nonsynonymous and nonsense) variants comparing the rapid decliners (N=130) and the slow decliners (N=119). Results: 161,774 variants were noted, of which 80,059 (49%) were novel. 25,743 annotated transcripts had at least one functional coding variant. Integration of the GWAS and sequence data revealed >99% concordance over ~7,500 variants sites. The four most significant genes for the CMC method collapsing variants with a minor allele frequency <5% included *CTH*, *GOLGA4* and *DCAF5* with protective effects and *HK2* with risk effects for lung function decline. The strongest signals under the Madsen-Browning weighted sum test (weighting by frequency in the slow-decliner group) included *ANKRD6, TRIM47, HLA-B* and again *HK2*. The number of variants for each test ranged from 3 (*CTH*) to 46 (*HLA-B*). We examined capture rates for these rare variants by the GWAS platform with a tagging approach. Conclusions: We identified several loci not detected in our GWAS that appear to harbor rare variants associated with lung function decline in COPD. While sampling only extreme subjects for sequencing optimizes statistical power to detect rare variants controlling lung function decline, we show that our genomewide marker panel did a poor job of tagging rare variants and an integrative approach of a GWAS in a larger sample (4,102) with the exome sequence in a smaller, highly selected sample (N=251) can be important in identifying genetic control of complex quantitative phenotypes.

You may contact the first author (during and after the meeting) at rmathias@jhmi.edu

---

Poster Session

10 a.m.-7 p.m.
Exhibit Hall, Level 2, Convention Center

2 p.m.-3 p.m.
Poster Category: Molecular Basis of Mendelian Disorders

Program Number 1129F

Mechanism and clinical severity of FIG4 mutations in Charcot-Marie-Tooth Disease.
M. H. Meisler, *et al*
Includes from NHGRI: L. G. Biesecker

Charcot-Marie-Tooth type 4J (CMT4J) is a severe recessive form of CMT caused by mutations of the lipid phosphatase FIG4. Eight new cases were identified through screening with the CMT gene panel. Three additional cases were found in a small cohort characterized by early onset and proximal as well as distal muscle weakness. These 11 previously unreported patients exhibited variable onset and severity, assymetric proximal as well as distal muscle weakness, EMG findings indicative of denervation in proximal and distal muscles, and frequent progression to severe amyotrophy and wheelchair dependence. Among the 16 known cases, fifteen have the compound heterozygous genotype FIG4I41T/null. The I41T mutation impairs interaction between the FIG4 protein and the associated scaffold protein VAC14 that is required for in vivo stability of the FIG4 protein, as demonstrated by the loss of FIG4 protein in mice homozygous for a null mutation of VAC14. FIG4-I41T is thus a hypomorphic allele encoding a protein that is unstable in vivo. A transgenic mouse model of CMT4J was generated by expression of the I41T mutant cDNA as

a transgene in Fig4 null mice. Expression of only 10% of the normal level of FIG4 protein was sufficient to rescue survival and prevent neurodegeneration in the null mouse. A mouse line with lower expression survives for 3 to 6 months and provides a model of CMT4J that can be used to test therapeutic intervention to increase the abundance of the I41T protein above the threshold for survival. The observations in the mouse model suggest that up-regulation of the I41T allele in CMT4J patients could be therapeutic.

You may contact the first author (during and after the meeting) at meislerm@umich.edu

Program Number 1139F

Clinical, Molecular, and Diagnostic Features of Hermansky-Pudlak Syndrome Type 3.

T. Gall, *et al*
Includes from NHGRI: R. Hess, T. Markello, K. O'Brien, H. Dorward, B. Gochuico, N. Cardillo, A. Cullinane, W. Gahl, M. Huizing

Hermansky-Pudlak syndrome (HPS) is a heterogeneous disorder of albinism, bleeding and sporadic other features including recurrent infections, granulomatous colitis and pulmonary fibrosis. There are 8 human HPS subtypes. All subtypes are caused by mutations in genes coding for proteins that are organized into four complexes, designated adaptor complex-3 (AP3; HPS subtype 2), Biogenesis of Lysosome-related Organelles Complex (BLOC)-3 (HPS subtypes 1 and 4), BLOC-2 (subtypes 3, 5, and 6), and BLOC-1 (subtypes 7, 8, and 9). After analyzing our NIH HPS cohort of 266 individuals with albinism and absent platelet dense bodies, HPS-3 was the most common subtype after HPS-1. More patients were affected with HPS-1 (195 patients; 148 with north-west Puerto-Rican 16-bp founder duplication, 47 with other mutations) than HPS-3 (34 patients; 9 with central Puerto-Rican 3,9-kb deletion, 25 with other mutations). And there were more unique mutations in *HPS1* (26) than in *HPS3* (23), but both these genes harbored significantly more mutations than other HPS genes: *HPS2* (4), *HPS4* (15), *HPS5* (14), *HPS6* (11), *HPS8* (1), *HPS9* (1). These findings classify HPS-3 as a major HPS subtype. We report 16 novel *HPS3* mutations, including missense, nonsense, frameshift, and splice site mutations, located throughout the *HPS3* gene. We show that BLOC-2 deficiency can be diagnosed by Western blotting of patients' cell extracts using HPS5 antibodies, assisting in subtyping HPS-3 patients. Clinical analysis of the 12 unreported HPS-3 patients carrying these mutations confirmed the relatively milder phenotype of HPS-3. In contrast to AP3 and BLOC-3 deficient patients, there are no cases of molecularly subtyped HPS-3 patients (or other BLOC-2 patients) with abnormal lung findings consistent with pulmonary fibrosis, death from respiratory failure, or need for lung transplant. However, HPS-3 patients are at risk, like all other HPS subtypes (15-20%), for an inflammatory bowel disease which includes granulomatous colitis distinctive from but as severe as Crohn's disease.

You may contact the first author (during and after the meeting) at trgall@carilionclinic.org

---

2 p.m.-3 p.m
Poster Category: Ethical, Legal, Social and Policy Issues in Genetics

Program Number 1355F

Are U.S. Doctors Receiving DTC Genetic Tests Results from Their Patients?

V. L. Bonham, *et al*
Includes from NHGRI: C. Clark, W. G. Feero

Background: Since 2005, the number of health-related direct-to-consumer (DTC) genetic tests has increased dramatically. Little is known about how often United States primary care physicians (PCPs) receive DTC genetic test reports and use the results in clinical care. **Methods:** We analyzed data from The Health Professionals' Genetics Education Needs Exploration Survey(HPGENE), a national survey of 787 general internists fielded from April to December 2010 (45% response rate), to examine the frequency with which PCPs received genetic test results in the preceding year, characteristics of PCPs who received these results, the extent to which DTC results were incorporated into clinical care, and the association of physicians' participatory decision-making (PDM) style with receipt of DTC results. **Results:** Of the 787 physicians that responded to the survey 505 were men and 479 were older than 45 years of age. Over half of the PCPs that received DTC test results worked in group or staff model practices (33%) and academic health centers (22%). Nineteen percent of respondents reported receiving DTC genetic test results from at least one patient in the preceding year. A higher proportion of female than male physicians reported seeing patients who presented DTC genetic test reports (24% vs. 16%, $p<0.05$). Physicians who had genetics training in residency (31%) and those who rated their knowledge of genetics as excellent, very good, or good (24%) were more likely to receive test results than physicians who had no training in residency (17%) and those who reported fair or poor knowledge (15 %), both p-values <0.01. Among physicians who received genetic test results, 59% discussed the results with the patients, 46% reviewed and placed the

results in the patients' medical records, 14% ordered additional tests, 10% incorporated the results into their treatment plans, and 7% referred the patient to a genetic specialist. Physicians' PDM style score was not associated with their use of patients' DTC test results in clinical care. **Conclusions:** A substantial minority of general internists are receiving patients' DTC genetic test results. Physician gender, practice setting, genetics training during residency, and knowledge of genetics are positively associated with receiving genetic test results. Importantly, physicians are incorporating this data into the medical record and clinical decision making.

You may contact the first author (during and after the meeting) at bonhamv@mail.nih.gov

---

2 p.m.-3 p.m.
Poster Category: Cardiovascular Genetics

Program Number 331F

Association of sequence variants in the USF1, ROS1 and ABCA genes with glycohemoglobin levels in the ClinSeq Study.

H. Sung1, *et al*
Includes from NHGRI: M. Krishnan, D. Ng, S. Gonsalves, P. Cruz, J. Mullikin, L. Biesecker, A. Wilson

ClinSeq is a large scale medical sequencing study designed to investigate associations of rare sequence variants to traits related to the risk of developing coronary heart disease (CHD). In this study, the samples are comprised of non-smoking patients, ages 45 to 65 with normal to severe coronary artery calcification. Forty-three phenotypes related to CHD, including glycohemoglobin level, were measured at the NIH Clinical Research Center in Bethesda, MD. Sanger-based sequencing of selected coding regions across 250 genes was performed at the NIH Intramural Sequencing Center. Each genotype was converted to missing if its call was ambiguous, if the Polyphred score was less than 99 or if either forward or reverse reads were missing. After removing sequence variants with a location-specific calling rate of less than 60% and individuals with a genotyping rate of less than 60%, a total of 8,837 sequence variants (SVs) were analyzed in 436 Caucasians. Of 8,837 SVs, 5,035 SVs were not observed in either dbSNP or the 1000 Genomes databases and were considered to be novel SVs. Tests of association of glycohemoglobin level (adjusted for age, sex and abdominal circumference) and each sequence variant were performed with simple univariate linear regression assuming additive allelic effects using PLINK v1.07. Thirty-four SVs, of which 27 were novel, with estimated minor allele frequencies of less than 0.003, were associated with both untransformed and transformed glycohemoglobin level at a significance level of less than 1e-08. In most cases the minor allele of these significant variants occurred only once in the entire sample. These findings will be investigated further by collapsing rare variants by position and by functional domain. The most significant SVs (p-values less than 1e-16) were all novel and included variants in the intronic regions of USF1 (chromosome 1 at 159,276,835 bp), ROS1 (chromosome 6 at 117,749,568 bp) and ABCA1 (chromosome 9 at 106,587,796 bp). USF1 is associated with antilipolytic insulin sensitivity and metabolic risk factors for cardiovascular disease, ROS1 is associated with myocardial infarct and ABCA1 is associated with atherosclerosis, CHD and cholesterol. These significant findings suggest the importance of the USF1, ROS1 and ABCA1 genes in determining glycohemoglobin levels and provide insight into the underlying genetic mechanisms of coronary heart disease in the ClinSeq study.

You may contact the first author (during and after the meeting) at sunghe@mail.nih.gov

---

2 p.m.-3 p.m.
Poster Category: Statistical Genetics and Genetic Epidemiology

Program Number 625F

A phenome-wide analysis of SNPs in the National Human Genome Research Institute genome-wide association catalog.

J. Denny, *et al*
Includes from NHGRI: R. Li, T. Manolio

Candidate gene and genome-wide association studies (GWAS) have provided greater understanding of the genetic underpinnings of many traits. Genetic data coupled to electronic medical records (EMR) offer the possibility of the inverse experiment, a "hypothesis-free" analysis of many phenotypes, or phenome-wide association scan (PheWAS). In this study, we used a cohort of 13,859 European-ancestry individuals genotyped in the Electronic Medical Records and Genomics (eMERGE) network. We performed PheWAS on 945 SNPs significantly associated ($p<5x10-8$) with traits in the NHGRI GWAS catalog. We examined each SNP for association with 947 phenotypes algorithmically defined using billing codes derived from normal clinical processes. We used logistic regression models for each phenotype adjusted for age, gender, and principal components.

PheWAS replicated 42%; (226/545) of GWAS catalog SNP-trait associations in 57 phenotypes, and 63% (82/130) of associations sufficiently powered for replication given our case size. A total of 46 SNP-phenotype associations in 24 phenotypes achieved genome-wide significance in PheWAS; 34 of these were previously known. Replications included Alzheimer's disease (*APOE*/*TOMM40*, p=1x10-25), hemochromatosis (*HFE*, p=2x10-25), type 2 diabetes (*TCF7L2*, p=3 x10-16), ischemic heart disease (9p21.3, p=1.10-12), and prostate cancer (8q24.21, p=9x10-8). New associations included acquired hypothyroidism (p=4x10-12) with *FOXE1* variants previously associated with thyroid cancer; skin neoplasms (p=4x10-17) and actinic keratosis (p=8x10-11) with *IRF4* variants previously associated with skin color; and aplastic anemia (p=3x10-6) with *WDR66* variants previously associated with mean platelet volume. Known pleiotropy at the 9p21.3 region (myocardial infarction, abdominal aortic aneurysm, and cerebrovascular disease) was replicated, and new associations were suggested (carotid stenosis, asthma, atherosclerosis, and other aneurysms, all p<10-5). *IRF4* variants were the most pleiotropic, with 14 phenotypes associated at p<10-5. EMR-based PheWAS replicated many SNP-disease associations in a real-world cohort, and thus can serve as a tool to improve confidence in GWAS associations. This analysis also suggested new, biologically plausible associations ("hypothesis-generating") for further study. Future use of PheWAS in tandem with traditional genetic analysis may help define pleiotropy and improve understanding of the genetic architecture of complex diseases.

You may contact the first author (during and after the meeting) at josh.denny@vanderbilt.edu

---

2 p.m.-3 p.m.
Poster Category: Clinical Genetics and Dysmorphology

Program Number 969F

UGT1A1 is a Major Locus Influencing Bilirubin Levels in African Americans.

G. Chen, *et al*
Includes from NHGRI: E. Ramos, A. Adeyemo, D. Shriner, J. Zhou, A. Doumatey, H. Huang, A. Bentley, H. Xu, B. Charles, C. Rotimi

BACKGROUND: Total serum bilirubin, an important byproduct of heme metabolism, has been associated with several clinical outcomes, including cardiovascular disease, diabetes and drug metabolism. Previous genome-wide association studies (GWAS) of serum bilirubin performed in European and East Asian populations reported that variants in the UGT1A1 gene (2q37.1) significantly influence serum total bilirubin levels. These findings have not been replicated in African ancestry populations; thus, we conducted a GWAS in African Americans to attempt to replicate reported findings and to identify novel loci influencing serum total bilirubin levels. METHODS: A total of 619 healthy unrelated individuals who participated in the Howard University Family Study (HUFS) were included in this study. Using a dense panel of over two million genotyped and imputed SNPs, we conducted a GWAS in African Americans for serum total bilirubin level. We assumed an additive genetic model and all statistical models were adjusted for age, sex, and the significant principal component from the sample covariance matrix of genotypes. RESULTS: Heritability for serum total bilirubim was estimated to be 0.42. Thirty-nine SNPs spanning a 78 kb region within the UGT1A1 gene (2q37.1) displayed p-values lower than the pre-determined genome wide significance threshold of 5×10-8. The lowest p-value was 1.7×10-22 for SNP rs887829. Notably, none of the other 38 SNPs in the UGT1A1 gene remained statistically significant in conditional association analyses that adjusted for SNP rs887829. We showed that rs887829 is in tight LD (r2 ≥ 0.74) with rs10929302 (-3156G > A, UGT1A1*91) located in the phenobarbital response enhancer module that is about 3kb upstream of the (TA)n variant reported to a better predictor of toxicity irinotecan - a cancer drug. We also replicated the reported association between variants in the SEMA3C and bilirubin in this cohort of African Americans. CONCLUSIONS: We observed that UGT1A1 is a major locus influencing total bilirubin levels in African Americans, adding to the evidence that this gene plays an important role in the determination of bilirubin level in human populations with different ancestral backgrounds. Our findings may also contribute to the understanding of the etiology of hyperbilirubinaemia and the pharmacogenomics role of UGT1A1 variants in drug (e.g., irinotecan) metabolism in African Ancestry populations.

You may contact the first author (during and after the meeting) at chengu@mail.nih.gov

Program Number 1034F

Long Term Survival in TARP Syndrome and Confirmation of *RBM10* as the Disease Causing Gene.

K. W. Gripp, *et al*
Includes from NHGRI: J. J. Johnston, C. Krause, L. G. Biesecker

TARP syndrome (talipes, ASD, Robin sequence, and persistence of left superior vena cava LSVC), is X-linked with pre- or postnatal lethality in males. Exon

resequencing in 2 families identified *RBM10* as disease causing gene. Clinical report: Pt. 1 was born preterm with U-shaped cleft palate, ASD and persistent LSVC, facial dysmorphia and cryptorchidism. He had severe respiratory distress and at age 3 ½ years remains ventilator-dependent at night and tube-fed. Brain anomalies include partial ACC, dysplastic caudate, and megacisterna magna. He has hypotonia, sensorineural deafness, optic atrophy and cortical visual impairment. Atrial flutter required ablation. Results: A novel *RBM10* frameshift variant was found in Pt. 1 and his mother was heterozygous.

Discussion: We confirm *RBM10* as pathogenic for TARP. This first TARP patient with long term survival expands the phenotype to include brain anomalies; hypotonia and developmental delay; deafness and optic atrophy; arrhythmia and characteristic facial features. Maternal heterozygosity underscores the importance of accurate diagnosis and counseling.

You may contact the first author (during and after the meeting) at kgripp@nemours.org

---

2 p.m.-3 p.m.
Poster Category: Therapy for Genetic Disorders

Program Number 1467F

Current treatment and outcome of infantile cobalamin C disease.
N. Carrillo-Carrasco, *et al*
Includes from NHGRI: J. Sloan, I. Manoli, N. Hauser, C. P. Venditti

Background: Cobalamin C disease (cblC), thought to be the most common inborn error of intracellular vitamin B12 metabolism, is caused by mutations in MMACHC and results in impaired synthesis of adenosyl- and methylcobalamin. Manifestations of this multisystemic disorder include impaired growth, neurologic and ophthalmologic abnormalities, and decreased lifespan. **Methods**: We evaluated 22 patients with infantile cblC disease (age range, 2 to 27 years) through NIH study 04-HG-0127 (clinicaltrials.gov: NCT00078078) to define their outcomes under current management strategies. **Results**: The median age of detection of patients diagnosed clinically was 2.6 months (range 0.6-24 months, n=16), compared to 0.3 months (n=3) by newborn screening. Common presentations included failure to thrive, encephalopathy, poor eye contact and nystagmus. All patients were treated after diagnosis, with considerable clinical improvement. We found that management was highly variable among patients. Protein restriction, although controversial, was seen in 41% of patients, with a median intake of 70% of their RDA. All patients had long-term complications despite treatment. Head circumference was the growth parameter most affected (median z-score -0.97; range -0.5 to -3.4). Seizures were reported in 55% and included refractory (14%; n=3), controlled (9%; n=2) and past history of seizures (32%; n=7); resolution coincided with lowering tHcy levels (n=4). The mean Vineland-II Adaptive Behavior Composite score was over 3 SD below normal (n=14, 52.9 ±18.87), but earlier age of initiation of treatment correlated with improved neurocognitive outcome (n=13, R=-0.755, p=0.003). Patients had progressive visual loss with a mean acuity of 20/500 (range 20/50 to 5/400). Fundoscopy revealed the characteristic findings of maculopathy, pigmentary retinopathy and/or optic pallor. Mean plasma MMA was 20 µM (range: 0.49-151µM), tHcy 70.8 µM (22-134 µM) and methionine 19.5 µM (4-53 µM). Most patients had trough plasma vitamin B12 levels below 1,000,000 pg/ml, a level noted to be therapeutic in other studies. **Conclusion**: We performed a single center study to define the outcome of infantile cblC. Early diagnosis and treatment seem to improve neurocognitive outcome, but not the progression of eye disease. Current treatment strategies are variable and commonly suboptimal. This study also provides baseline information to evaluate the efficacy of newborn screening and novel therapeutic interventions.

You may contact the first author (during and after the meeting) at carrilln@mail.nih.gov

Program Number 1481F

Gene therapy as a potential treatment of cardiomyopathy in propionic acidemia.
R. J. Chandler, *et al*
Includes from NHGRI: N. Carrillo-Carrasco, S. Chandrasekaran, C. P. Venditti

Propionic Acidemia (PA) is autosomal recessive metabolic disorder caused by a deficit in the enzymatic activity of propionyl-CoA carboxylase. Genetic defects in *PCCA* or *PCCB*, which code for propionyl-CoA carboxylase, result in PA. Patients with PA exhibit elevated levels of propionic acid and methylcitrate in their plasma and urine, and can present with potentially lethal metabolic decompensations, hyperammonemia, neurological complications, and a potentially lethal cardiomyopathy. These medical problems can arise and the cardiomypathy does progress in spite of current medical treatments. A therapy that stabilized the patients and/or addresses the progressive and often lethal cardiomypathy would be beneficial to patients. We have previously demonstrated the ability of rAAV8 gene delivery to rescue the murine model of PA from lethality and decrease the levels of

disease related metabolites. In light of the cardiomyopathy observed in the human disease, the cardiac histology of the PA mouse model was examined to determine whether cardiac abnormalities were present. Oil red staining revealed an abnormal accumulation of lipid and electron micrographs (EMs) revealed gross abnormalities of the cardiac tissue in the PA mice. Specifically, the cardiac myofibers were poorly organized and reduced in number. However, cardiac EMs from the PA mice treated by rAAV8 gene delivery of the *PCCA* gene exhibited no abnormal histopathological changes. These results demonstrate the potential efficacy of rAAV8 gene delivery as an alternative treatment to transplantation for the potentially lethal cardiomyopathy observed in PA.

You may contact the first author (during and after the meeting) at rchandler@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Molecular Basis of Mendelian Disorders

Program Number 1124F

A germline mutation in a family with Gaucher disease.
N. Tayebi, *et al*
Includes from NHGRI: H. Saranjam1, E. Sidransky

Gaucher disease (GD), the most common lysosomal storage disorder, results from the inherited deficiency of the enzyme glucocerebrosidase (GCase). Over 300 unique mutations have been identified in GBA, the gene encoding glucocerebrosidase. GD is autosomal recessively inherited, so affected individuals inherit mutations from both parents. We investigated a family with a baby diagnosed with type 2 GD. Direct sequencing of GBA demonstrated that the child's genotype was T323I /L444P. We also sequenced parental DNA. Mutation T323I was inherited from the father, but no mutation was detected in a DNA sample from the mother. This was confirmed by sequencing GBA in DNA from different maternal cells, including blood, skin fibroblasts, and buccal cells. Further analysis demonstrated that the mother's somatic cells had normal levels of GCase expression and enzymatic activity. Studies of 32 additional markers evaluated in the child and parents confirmed maternity. Based on the experiments performed, we propose that the L444P mutation is likely present only in the mother's germ line cells, and was passed on to the proband. This is the first report of the presence of a germ line mutation in Gaucher disease, and can have significant implications for genetic counseling in recessive disorders.

You may contact the first author (during and after the meeting) at tayebin@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Therapy for Genetic Disorders

Program Number 1466F

Cell-based evaluation of small molecules for treatment of Pompe disease.
W. Westbroek, *et al*
Includes from NHGRI: A. M. Gustafson, J. J. Marugan, J. Xiao, W. Zheng, A. Velayati, E. Goldin, E. Sidransky

Glycogen storage disease II or Pompe disease is an autosomal recessive lysosomal storage disorder (LSD) caused by dysfunction of the lysosomal enzyme acid alpha-glucosidase (GAA). Mutations in GAA impair the breakdown of glycogen, causing storage of the substrate in enlarged lysosomes. Cardiac and skeletal muscles are severely affected, causing impaired breathing and mobility. More than 100 distinct disease-causing mutations in GAA have been identified, some of which retain enzymatic activity in vitro, but are not trafficked to the lysosome. Instead the misfolded GAA stays in the endoplasmic reticulum (ER) and eventually undergoes proteosome-mediated breakdown. Enzyme replacement therapy with Myozyme ®, the only current treatment for Pompe disease, improves clinical outcome, but the treatment remains expensive, inconvenient and does not reverse all disease manifestations. It was postulated that small molecules which aid in protein folding, ER trafficking, and translocation of enzymes to lysosomes could provide an alternate therapeutic strategy. We identified several promising non-inhibitory compounds for GAA from a chemical compound library by performing distinct in vitro assays. For further biological evaluation of these GAA compounds, we developed an in vivo cell-based immune-fluorescence cytochemistry assay to assess chaperone capacity. Translocation of GAA to the lysosomes in Pompe cells was evaluated by measuring the extent of co-localization between GAA and Cathepsin-D, a lysosomal marker, using laser scanning confocal microscopy. Using a GAA-specific antibody, control fibroblasts showed specific GAA staining in the lysosomes, while fibroblasts from patients with Pompe disease lacked this signal with immune-fluorescence cytochemistry. However, after six days of treatment of Pompe fibroblasts by adding the GAA activator to the growth medium, the expression level and translocation of GAA to lysosomes was comparable to control cells. Such activators are particularly desirable chaperones, since enzyme function remains intact. These promising small molecules, identified using an in vivo cell-based assay,

merit further evaluation as a potential new therapy for Pompe disease.

You may contact the first author (during and after the meeting) at wwestbro@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Complex Traits: Theory and Methods

Program Number 418F

Substance abuse and addiction measures in the PhenX Toolkit for use in genomic and epidemiologic studies.
K. Conway, *et al*
Includes from NHGRI: H. A. Junkins, E. M. Ramos

The PhenX (consensus measures for **Phen**otypes and e**X**posures) Toolkit (https://www.phenxtoolkit.org/) offers 295 high-quality, well-established, standard measures of phenotypes and exposures for use in genome-wide association studies and other large-scale genomic and epidemiologic studies. Currently, the Toolkit addresses 21 research domains (fields of research), including an Alcohol, Tobacco and Other Substances domain with 14 measures related to substance abuse and addiction (SAA) (e.g., cigarettes per day). There was a recognized need for additional SAA measures in the Toolkit, so a project was launched in early 2011 to create six "Specialty" Collections and one Core" Collection of SAA-related measures. These new Collections will provide SAA and other investigators with common measures, and thus make it easier to compare or combine studies. It is desirable to replicate association findings by comparing results from different studies or to combine studies to create larger sample sizes, increasing statistical power and the ability to detect more subtle and complex associations. To provide guidance, a SAA Scientific Panel (SSP) of 10 academic and federal governmental scientists was assembled to select and define the initial scope of six specialty areas using a consensus process. The six specialty areas are: (1) Assessment of substance use and substance use disorders, (2) Substance-specific intermediate phenotypes, (3) Substance use-related neurobehavioral and cognitive risk factors, (4) Substance use-related psychosocial risk factors, (5) Substance use-related community factors, and (6) Substance use-related co-morbidities and health-related outcomes. Each of three Working Groups (WGs) composed of academic and government experts will address two Specialty Areas. The WGs will recommend up to eight high-priority measures for each Specialty Area. The WGs will use a consensus-based process that includes input from the scientific community as they select the measures for inclusion in the PhenX Toolkit. Outreach results will be reviewed by the WGs and considered in their final deliberations. The WGs present their selections for review and approval by the SSP and the PhenX Steering Committee. We present the rationale, criteria, scope, and preliminary results from the SSP and the SAA Working Groups. It is expected that some 48 SAA measures will be added to the Toolkit by spring 2012. *Supported by NHGRI and NIDA grants 5U01HG004597 and 3U01HG004597-03S3.*.

You may contact the first author (during and after the meeting) at Kevin.Conway@nih.gov

Program Number 432F

Genome-wide analysis of imputed genotypes identifies chemokine receptor-1 (*CCR1*) as a novel susceptibility locus in Behçet's disease. *E. F. Remmers1, G. Bertsias1, 2, Y. Kirino1, M.* J. Ombrello
Includes from NHGRI: C. Satorius, D. L. Kastner

Behçet's disease (BD) is a genetically complex disease characterized by recurrent inflammatory attacks affecting the orogenital mucosa, eyes, and skin. We have previously performed a genome-wide association study with 311,459 SNPs in 1,215 BD cases and 1,278 healthy controls from Turkey, and have identified independent associations with HLA-B*51, an additional MHC class I locus, *IL10*, and the *IL23R-IL12RB2* locus. In this study, we carried out whole-genome imputation, using as a reference, the genotypes of 96 of the Turkish healthy controls determined on Illumina HumanOmni1M-Quad SNP chips. SNPs were excluded for deviation from HWE ($p<5\times10\text{-}4$), low call rate (<95%), and low MAF (<5%). Imputation was conducted using MACH v1.0.15 providing 814,474 SNPs for analysis in the 1,215 BD cases and 1,278 controls. Using a p-value cut-off of $1\times10\text{-}5$, we identified 114 non-HLA gene SNPs suggestive of association with BD. One imputed SNP rs7616215 on chromosome 3, located ~38 kb from the 3' UTR of the chemokine receptor-1 gene (*CCR1*), (odds ratio [OR] = 0.71, $p=1.9\times10\text{-}8$) exceeded genome-wide significance ($p<5\times10\text{-}8$). Fine mapping of the *CCR1* region using Sequenom iPLEX assays confirmed the imputation results for rs7616215 and identified 2 additional SNPs in strong LD with rs7616215 that also exceeded genome-wide significance. The association of rs7616215 replicated in additional Turkish and Japanese BD cases and controls (in a meta-analysis of 2,007 cases and 2,187 controls, OR = 0.73, 95% CI 0.66-0.81, $p=3.1\times10\text{-}10$). CCR1 belongs to the family of CC-motif chemokine receptors, is expressed on neutrophils, monocytes, and T lymphocytes, and binds several chemokine ligands, including CCL5/RANTES, CCL3/MIP-1 α, and CCL4/MIP-1 β. A role for CCR1 has been identified in several inflammatory conditions, such

as rheumatoid arthritis, multiple sclerosis, and transplant rejection. ENCODE data indicate that rs7616215 resides in a putative regulatory genomic domain. Analysis of *CCR1* transcripts from the HapMap European (CEU), Chinese (CHB), and Japanese (JPT) subjects showed the disease-associated (T) allele correlated with significantly reduced *CCR1* expression ($p<0.03$). Similarly, we found reduced *CCR1* mRNA as a function of the number of rs7616215 risk alleles in purified peripheral blood CD14+ monocytes from healthy controls. We have identified *CCR1* as a novel susceptibility locus in BD, with potential implications for regulation of inflammatory responses in the context of disease.

You may contact the first author (during and after the meeting) at remmerse@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Psychiatric Genetics, Neurogenetics and Neurodegeneration

Program Number 550F

The pervasive neurologic phenotype of mild Chediak-Higashi Syndrome.
C. Toro, *et al*
Includes from NHGRI: A. Cullinane, W. Westbroek, M. Huizing, C. Groden, W. Gahl, W. Introne

Chediak-Higashi syndrome (CHS) is a very rare autosomal recessive condition due to mutations in lysosomal trafficking regulator gene (*LYST/CHS1* gene, OMIM #606897). The CHS1 protein is thought to function in the formation and trafficking of lysosome-related organelles. Clinically, CHS is characterized by partial albinism, a bleeding diathesis, recurrent infections that are often ultimately fatal, and hemophagocytic lymphohistiocytosis, or the accelerated phase. Giant inclusions within white blood cells are diagnostic. Mortality is high in the first decade of life, and the only available treatment is bone marrow transplantation (BMT). Some CHS patients present with mild findings, i.e., inconspicuous hair and eye hypopigmentation and subtle cognitive difficulties. By their early 20s, these patients manifest slowly progressive neurological deterioration, resembling adults with typical CHS who underwent successful BMT as children. 8 young adults with confirmed CHS, but without BMT, were evaluated at the National Institutes of Health Clinical Center. Testing included neurologic examination, neuropsychiatric testing, EMG, and MRI of the brain. The CHS neurological phenotype involves a mixture progressive length-dependent peripheral neuropathy with foot drop, distal atrophy, weakness, sensory loss and Parkinsonism. Cognitive function, only mildly impaired in childhood, declines progressively with worsening executive dysfunction. Cerebellar signs, basal ganglia dysfunction, upper motor neuron and posterior column dysfunction develop progressively. By their 3rd decade, most patients have diminished capacity for activities of daily living. The pervasive, late neurologic features of CHS, which occur despite BMT, are under-recognized and can appear without the classical features of the disease. The length dependent nature of several neurological features in CHS suggests a possible disruption of axonal transport within neurons. CHS should be considered in the differential diagnosis of patients with unexplained peripheral neuropathy, cognitive decline, cerebellar symptoms, Parkinsonism or long tract findings.

You may contact the first author (during and after the meeting) at toroc@mail.nih.gov

---

3 p.m.-4 p.m.
Poster Category: Statistical Genetics and Genetic Epidemiology

Program Number 612F

Analyze X: A Comparison of X Chromosome and Autosomal Results in GWAS.
A. Wise, *et al*
Includes from NHGRI: L. Gyi, T. Manolio

Genome-wide association studies (GWAS) have identified over 1,200 associations with $p<5x10-8$ for over 200 traits, yet only 7 such associations have been reported on the X chromosome (chr). This raises the question: why is the X chr so poorly represented in GWAS results? A review of over 300 GWAS suggests that this is largely due to exclusion of X chr variants from analyses as one of the first steps of quality control, despite these regions being assayed on many current microarray platforms. Reviewing the NHGRI Catalog of Published GWAS (www.genome.gov/gwastudies) showed that only 121 of 374 (32%) of GWAS published from Jan 2010 through Mar 2011 reported analyzing the X chr in their Methods sections. Accounting for the number of genes on the X chr also does not eliminate the gap in GWAS associations; as the X chr (1,669 genes) has only 7 reported associations at 5 distinct loci, while chr 7 (1,880 genes) has 48 reported associations at more than 20 distinct loci. Interrogating all of the variants found in papers that identified X chr hits at $p<1x10-5$ in the NHGRI Catalog, to account for sample size and genotyping platform variations, X chr variants have similar minor allele frequencies (MAFs) as autosomal variants (X chr avg

MAF=0.34 and autosomal avg MAF=0.39, p=0.27). Median p-values for the hits found on the X chr were higher, however, by approximately 1 order of magnitude than those seen for autosomal variants in these same studies (p=0.04), indicating that power may indeed be an issue perhaps due to smaller effective chr numbers. Comparing the functional classes assigned to these variants by NCBI, none (0/51) of the variants on the X chr were missense or synonymous variants, as compared to 5% (27/582) of the autosomal variants. Intronic variants were found in similar proportions in the X chr and autosomal hits, 31% and 37% respectively; while intergenic variants, on the other hand, were more common amongst the X chr hits, 67% versus 49%, p=0.02. Given the similarity found between the X chr and other autosomes in variant MAFs, and overall number of genes, these factors cannot explain why the X chr is so poorly represented in GWAS results. However, the higher p-values found on the X chr and lack of missense or synonymous variants point towards potential issues with power. Improvements in genotype calling accuracy and methods developed specifically for the X chr may thus improve power to detect important associations on the X chr in future GWAS.

You may contact the first author (during and after the meeting) at anastasia.wise@nih.gov

---

3 p.m.-4 p.m.
Poster Category: Genomics

Program Number 814F

VAR-MD: Advanced analysis of whole exome sequencing for detection of hemizygosity. M. Sincan, *et al*
Includes from NHGRI: D. R. Simeonov, T. C. Markello, D. A. Adams, C. Toro, C. Tifft, W. A. Gahl, C. F. Boerkoel

Whole Exome Sequencing (WES) is increasingly used in diagnoses of rare Mendelian diseases, but methods are required to comprehensively analyze and prioritize huge numbers of variants. We have developed a tool that can evaluate WES variants and produce a ranked list of potential causative mutations in the family. This tool 1) employs a simple, tab-delimited input file containing the annotated variants shared within the family; 2) utilizes family structure and affected status, 3) makes use of public data sets for further annotation and analysis; 4) makes use of the short read alignments of the exomes in bam format for additional analyses; and 5) for the first time, detects hemizygosity. This tool was highly successful in prioritizing and ultimately identifying pathological mutations causing spinocerebellar ataxia type 28, GM1-gangliosidosis, and spastic paraplegia type 35 (SPG35) in different patients enrolled in the NIH Undiagnosed Diseases Program. Of particular interest was the case of a boy with SPG35. Previous WES analytical tools read the mother's genotype as homozygous for the normal FA2H allele while the father was found to carry a point mutation. However, the maternal allele actually carried a deletion of exons 3-7. Since monoallelic deletions cause continuous loss of heterozygosity across the deletion, we interrogated the B a.m. alignment files of the family for unusual runs of homozygosity. We confirmed that this was indeed the case for this family, and incorporated this process into our WES filtering analysis. This advanced tool for WES analysis (VAR-MD), which has broad applicability and includes the innovation of detecting copy number variants, is available to the public at http://research.nhgri.nih.gov/.

You may contact the first author (during and after the meeting) at sincanm@mail.nih.gov

DAY 5
SATURDAY
OCTOBER 15, 2011

Platform Talks

1:30-3:30 p.m.
Platform Session 70 on Advances in Technology
Room 517D, Level 5, Convention Center

1:30-1:45 p.m.
Program Number 261

Indexing and deep sequencing of a point mutation in mosaic samples from Proteus syndrome: Alternative detection strategies.
M. J. Lindhurst, *et al*
Includes from NHGRI: J. K. Teer, J. J. Johnston, E. M. Finn, J. C. Sapp, J. C. Mullikin, L. G. Biesecker.

Detection of mutations in mosaic samples is challenging due to the variability in the level of the mutant allele that can occur in various tissues. Recently, we have shown that Proteus syndrome is caused by a somatic mutation that results in constitutive activation of a key signaling pathway in affected tissues. The variant was initially identified using next gen sequencing of 17 samples. To confirm and extend these findings, we screened 158 samples by Sanger sequencing. Seventy-one of these samples had clear evidence of the mutant allele on the electropherograms, 17 had electropherograms that suggested a low level of the mutation, and 70 were apparently negative. However, it was challenging to distinguish low levels (<1%) of mosaicism from background signals or noise in the electropherogram traces. We next developed a PCR/restriction endonuclease assay (MboII) that could be separated on the ABI3130 and was designed such that MboII digested only the mutant amplicon. Areas under the curve for the mutant and wild-type peaks on the instrument were assessed quantitatively. A validation experiment using dilutions of cloned mutant and wild-type DNA showed a correlation of r2=0.9993 with a lower limit of 1% mutant allele sensitivity. Using this assay, the mutant allele was detected in all 71 of the positive samples, 15 of the above 17 samples that were suggestive and 15 samples of the 70 samples that were negative by Sanger. This increased the number of patients with at least one positive sample from 23 to 27. The level of mutation in these samples ranged from 1% to ~50%. We screened 75 samples from non-Proteus individuals; all were negative for the mutation. While the PCR/MboII assay increased sensitivity compared to sequencing, variation increased at low mutant allele levels (again, <1%). To address this, we PCR-amplified the mutation from 12 samples using indexing primers and used next-gen sequencing to generate ultra-deep coverage. These samples were pooled, and sequenced on a single GAIIx lane (76 bp paired-end). 28,568,780 total reads were generated, and 335,378-1,649,523 reads were assigned to each sample. Allele frequencies were calculated at the variant position, and agreed well with the MboII results. Additional optimizations will be performed and have the potential to further improve sensitivity. (The authors confirm that the gene and mutation will be disclosed at the ICHG meeting should the abstract be selected for presentation).

You may contact the first author (during and after the meeting) at marjr@mail.nih.gov

1:30-3:30 p.m.
Platform Session 76 on Genetic Screening and Education
Room 710B, Level 7, Convention Center

3-3:15 p.m.
Program Number 315

Tryptic Peptide Analysis of WBC to Diagnose Genetic disorders: Application to Primary Immunodeficiency Disorders and Nephropathic Cystinosis.
S. Hahn *et al*
Includes from NHGRI T. Vilboux, W. Gahl

PURPOSE: There are many genetic diseases that would benefit from early diagnosis but that have no population screening platform. Examples include primary immunodeficiency disorders such as X-linked agammaglobulinemia (XLA), Wiskott Aldrich Syndrome (WAS), and Severe Combined Immune Deficiency (SCID), as well as nephropathic cystinosis. BTK, WASP, CD3 ε and cystinosin are low abundance proteins deficient in these conditions. They are localized in cell membranes, cytoplasm, or on the cell surface and are not detectable in plasma. We propose that these proteins can be observed in proteolytically digested extracts of WBCs and will be absent or significantly reduced in affected cells, making diagnosis possible. **METHODS:** Candidate peptides were screened by in silico trypsin digestion modeling followed by a BLAST search to insure that the sequences are unique within the human genome. The final signature" peptides were selected by evaluating the MRM chromatogram for the isotopically labeled peptide and the WBC digest peptide in control (n=20) and patient samples. Five lymphocyte cell lines were used to establish the absence of signature peptides for primary immunodeficiencies. Three PBMC samples from XLA were blindly analyzed. WBCs from 10 cystinosis patients with known genotypes were also tested. The amount of each

peptide in the WBC was determined by taking the ratio of the peak area for the signature peptide to that of the labeled peptide and reported as normalized to actin. **RESULTS:** Three blinded samples lacked only BTK peptides, confirming that these samples were from XLA patients. All five cell lines clearly showed the biochemical phenotype of each cell line. Cystinosin was non-detectable in 9 cystinosis patients. One patient showed a cystinosin peak but at a very low concentration compared to control. **CONCLUSION:** Our method quantified the proteolytic peptides for the target proteins, BTK, WASP, CD3 ε and cystinosin in various cell lines and patient samples. Targeted proteins from these conditions were either absent or significantly diminished. This approach can be potentially utilized as part of a multiplex analysis for many genetic conditions in which the protein of interest is significantly reduced. While the transition from white blood cells to dried blood spots will be challenging, we believe that with further enrichment and optimization, this method can be potentially applied to newborn screening.

You may contact the first author (during and after the meeting) at sihahn@u.washington.edu