## GENETICS

### G01. WITHDRAWN

### G02. ClinVar: A Centralized Database for Interpretations of Both Germline and Somatic Variants

*M. Landrum, J.M. Lee, M. Benson, G. Brown, C. Chen, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, W. Jang, K. Katz, M. Ovetsky, G. Riley, A. Sethi, R. Tully, W. Rubinstein, D. Maglott*
*National Institutes of Health, Bethesda, MD.*

**Introduction:** ClinVar is a public archive of variants and interpretations of their clinical significance. Interpretations are provided by clinical testing laboratories, research laboratories, locus-specific databases, OMIM and GeneReviews, as well as expert panels and practice guidelines. ClinVar has a broad scope, including sequence and structural variants; simple and complex alleles, chromosomal and mitochondrial; and both germline and somatic origins. **Methods**: ClinVar currently has interpretations for more than 2000 somatic variants in more than 850 genes, including BRCA1, BRCA2, EGFR, RB1, PIK3CA, KRAS, BRAF, FGFR3, and TP53. Interpretations for somatic variants are specific to the type of tumor in which the variant was observed and interpreted; therefore a variant may have multiple interpretations for different tumor types. The interpretation may indicate whether the variant contributes to pathogenicity (tumorigenicity or disease progression) or it may indicate whether the variant affects an individual's response to a drug. Evidence for the interpretation may include observations of individuals or groups of individuals with the variant. Observations of individuals may include age, sex, ethnicity; fields specific to somatic variants will be added based on feedback from the community. Evidence for an interpretation may also include literature citations or experimental assays that describe the functional significance of the variant. **Results**: Interpretations for the same variant are aggregated in ClinVar so that consensus or conflict is apparent, and submitted interpretations and evidence may be viewed together. A search for somatic variants in ClinVar may use a gene symbol, a variant name, the name of a tumor type, or a drug name. The search results page includes filters which can be used to narrow down the results; for example, results can be filtered to include only somatic variants reported from clinical testing. **Conclusions:** ClinVar welcomes additional submissions of any type of variant through the ClinVar Submission Portal (https://submit.ncbi.nlm.nih.gov/clinvar/). Once the submitting organization is registered and approved by NCBI staff, submitters can upload data in the submission portal. A wizard is available for those with a single interpreted variant to submit. The wizard guides the submitter through the process of describing the variant, the condition (either tumor type or drug response), the interpretation, and the observations provided as evidence. Submitters with more than one variant interpretation to submit fill out the submission spreadsheet templates which are then uploaded to the submission portal. More information about the submission process is available online (http://www.ncbi.nlm.nih.gov/clinvar/docs/submit/).

### G03. Technical Validation of a Next-Generation Sequencing Assay for Detecting Variants in Genes that Contribute to Inherited Cancers

*W. Xu[1], W. Xu[1], C.L. Sims[2], T. Cavness[1], N. Lee[1], A. Goldman[1], M. Ta[2], S. Covic[1], K. Hamada[1], C. Ambrosius[1], S.R. Opie[1]*
*[1]Vantari Genetics, Irvine, CA; [2]ResearchDx, Irvine, CA.*

**Introduction**: New cancer cases in the United States are estimated to reach almost 1.7 million in 2016. Sequence variations in germline DNA may contribute to increased risk. Targeted next-generation sequencing (NGS) gene panels are becoming increasingly accepted as an efficient screening tool to examine the genetic profiles of patients with a relevant clinical history. We describe the analytical validation of a 144 gene Human Hereditary Cancer Panel (HCP) which corresponds to over 90 types of cancers. This validation was established using College of American College Pathologists (CAP) and CLIA requirements to detect germline single nucleotide variations (SNV) and short insertions and deletions (indels). **Methods**: Genomic DNA was extracted from 24 well-characterized cell lines (Coriell Insitute for Medical Research, NJ) and 10 matched blood and saliva specimens from the same consented and de-identified donors. Qiagen Gene Read Library Preparation Kits were used to prepare PCR amplicon libraries for subsequent barcoding and sequencing. An aliquot of DNA extracted from blood was sent to an external CLIA/CAP accredited laboratory for whole exome sequencing. All sequencing was performed on the Illumina MiSeq platform using V3 chemistry. De-multiplexed FASTQ files were processed by a bioinformatics pipeline developed with Galaxy. Variant filtering and clinical annotation was conducted with commercially available software (Cartagenia, MA). Sequencing artifacts were identified by reviewing known artifacts from the Platinum Genome Sequence (Illumina) for NA12878. Errors due to base repeats, low coverage and low quality reference data were excluded after manual review. A total of 353 known variants and 12,176,970 wild-type bases were assessed for concordance by detection in cell lines and by comparing whole exome sequencing results in blood and saliva. **Results:** This assay detects SNV and small indel (<10bp) variants with an average depth coverage (DOC) of 147x and minimum DOC of 10x. Assay quality metrics for all combined variant types using unfiltered .vcf files were: accuracy, 99.81%; concordance, 98.81%; analytical specificity, 99.99%; analytical sensitivity 98.33%, positive predictive value, 98.60%, negative predictive value, 99.99%; inter-assay precision, 95.54%; intra-assay precision, 95.00%. Linear regression analysis (r^2=0.992) to calculate mean allele frequency for heterozygous calls was 0.46 (0.69 to 0.24). **Conclusion**: These results validate for clinical use a targeted NGS sequencing panel analysis for identifying mutations in 144 genes related to over 90 different inherited cancers.

### G04. A High-Throughput, Multiplex Digital Droplet PCR Method for the Detection of Severe Combined Immunodeficiency and Early T Cell Lymphopenia by Newborn Screening

*N. Vidal-Folch, D. Milosevic, R. Majumdar, D.K. Gavrilov, D. Matern, K. Raymond, P. Rinaldo, S. Tortorelli, R.S. Abraham, D. Oglesbee*
*Mayo Clinic, Rochester, MN.*

**Introduction:** Severe Combined Immunodeficiency (SCID) is an immunological condition with neatly universal mortality without early intervention via hematopoietic cell transplantation (HCT), gene therapy or enzyme replacement therapy (ERT). Infants with SCID have defective T-cell development resulting in significant T-cell lymphopenia (TCL). T-cell receptor excision circles (TREC) levels are widely used by newborn screening (NBS) programs for SCID. In most laboratories, qPCR is used to quantify TRECs in dried blood spots (DBS). Although effective, qPCR requires normalization to controls and is susceptible to variation due to DNA extraction efficiency. Here, we describe a novel multiplex, digital droplet PCR (ddPCR) method for TREC measurement in DBS and show a potential reduction in false positive rate for SCID. **Methods:** DNA from a 3.2mm DBS punch was extracted with a multistep method and denaturation at 99°C. Simultaneous quantification of TREC and RNAseP (housekeeping gene) was performed using AutoDG and QX200 ddPCR Bio-Rad system. A total of 649 infants were tested including 600 full-term infants, 10 preterm infants, 29 lymphocyte profiled infants and 10 clinically diagnosed infants (1 X-linked SCID, 1 RAG1 Omenn syndrome, 2 idiopathic T-cell lymphopenia (iTCL), 1 chr22q11.2 deletion negative TCL, 2 DiGeorge syndrome, 1 CHARGE syndrome, 1 ataxia telangiectasia and 1 cartilage hair hypoplasia). Results were reported as TREC copies/µl blood. **Results:** DBS TREC content for full-term infants was 14 to 474 TREC copies/µl blood, median of 128. Preterm infants ranged 28 to 186, median of 104. Patients diagnosed with SCID and other TCL conditions were confirmed to contain ≤15 TRECs. Clinical threshold was established at 20 TRECs, which is above the TREC content of diagnosed cases seen to date and reduces the false-positive rate, retest, and referral rate. The LLOQ was 14 TRECs. LOD was 11 TRECs. Precision experiments showed <20 %CV for intra-assay (at 54 TRECs) and inter-assay (at 60 TRECs) imprecision. Testing 29 DBS with known lymphocyte profiles indicated a clinical sensitivity of 88.8% and clinical specificity of 100%. **Conclusions:** We developed a highly sensitive and accurate multiplexed ddPCR method for absolute quantification of TREC in DBS without a standard curve. This method can detect SCID and other TCL conditions associated with absent or low TREC content. The method is cost-effective and suitable for NBS testing.

### G05. The Association Between Pro-Fibrotic Genes Over Expression and Epicardial Adipose Tissue Mass Increase in CAD Patients Can Be a Potential Powerful Diagnostic Tool in the Prevention of Cardiac Fibrosis

*M.M. Corsi Romanelli[1], E. Dozio[2], A. Sigrüner[3], A. Parolari[1], G. Schmitz[3], E. Vianello[2]*
*[1]University of Milan and IRCCS Policlinico San Donato of Milan, Milan, Italy; [2]University of Milan, Milan, Italy; [3]University Regensburg, Regensburg, Germany.*

**Introduction**: The main second messenger in the heart is cAMP and the over expression of its signaling can promote cardiac hypertrophy and fibrosis. During cardiac pressure overload that occurs in coronary artery diseases (CAD), the up-regulation of cAMP gene can directly over express Epac (exchange protein directly activated by cAMP) genes to amplify cardiac remodeling signaling. Epicardial adipose tissue (EAT) can directly influence myocardium because, sharing with it the same microcirculation, it can release its gene products into their common vessels. Farther an increase of EAT mass is now associated to the up-regulation of several genes involved in cardiac remodeling. Since cAMP has a pivotal role in the regulation of adipose stromal cells our aim is to elucidate if an increase of EAT mass in CAD and non-CAD patients can promote cardiac fibrosis signaling through Epac gene over expression. **Methods**: Twenty-three CAD patients underwent to bypass surgery and 10 patients underwent to valvular replacement (non-CAD) were enrolled at IRCCS Policlinico San Donato of Milan. EAT biopsies were collected after surgery in all protect reagent (Quiagen) and preserved –20°C until analysis. EAT thickness is examined by echocardiography using an M-mode color-Doppler VSF (General Electric) with a 2.5 to 3.5 MHz transducer probe. All patient were than stratified according to EAT median thickness value (8mm). *ADPY2* encoding for andenylate cyclase deputed to produce cAMP, and *RAPGEF5* and *RAPGEF6* genes from Epacs family are analyzed using one-color microarray platform (Agilent). Fifty ng of total RNA was labeled with Cy3 using the Agilent Low Input Quick-Amp Labeling Kit-1 color. cRNA was purified with the RNeasyMini Kit (Qiagen) and the amount and labeling efficiency were measured with NanoDrop. Hybridization was done using the Agilent Gene Expression Hybridization Kit and scanning with the Agilent G2565CA Microarray Scanner System. **Results**: CAD patients expresses higher ADPY2 mRNA than non CAD patients but no significant differences were found among groups classified according to EAT median thickness. CAD patients with EAT >8 mm presented higher RAPGEF5 expression level than non-CAD patients with EAT >8

mm; no statistical differences were found intra-groups. RAPGEF6 is higher expressed in CAD with EAT >8 mm than both non-CAD patients and in CAD with EAT<8mm. **Conclusion**: Our data show that in EAT of CAD patients, ADPY2 gene is up-regulated suggesting an over production of cAMP. This difference can be associated to pro-fibrotic genes up-regulation part of Epacs family, including RAPGEF5 and RAPGEF6, in response to EAT thickness increase. Therefore EAT measurement can be propounded as potential diagnostic tool in the prevention of cardiac fibrosis in CAD.

**G06. A Two-for-One Deal: Detecting Germline Cancer Predisposition Variants Concurrently with Tumor Somatic Mutations in 777 Patients Analyzed by MSK-IMPACT**

*D. Mandelker, Y. Li, M. Prasad, A. Syed, A.R. Balakrishnan, R. Chandramohan, D. Delair, S. Argueta, C. Yang, A. Arnold, M. Walsh, Z. Stadler, M. Robson, K. Offit, M. Berger, A. Zehir, M. Ladanyi, L. Zhang*
*Memorial Sloan Kettering Cancer Center, New York, NY.*

**Introduction:** To definitively identify somatic variants in clinical NGS tumor sequencing, matched normal DNA may also be sequenced, and germline variant calls subtracted from the tumor variant calls. Here, we examined the proportion of cancer patients presenting for somatic sequencing who harbor pathogenic germline variants for an undiagnosed hereditary cancer syndrome. **Methods:** Patients had both their tumor and normal DNA's sequenced for MSK-IMPACT, a 410 gene NGS panel, for the purpose of somatic tumor profiling. The patients were subsequently given the option of consenting to germline analysis for 76 genes associated with hereditary cancer predisposition. Germline variants were called in the normal DNA sequenced for MSK-IMPACT tumor testing and classified based on ACMG criteria. For all cases with a pathogenic or likely pathogenic germline variant, the patient's tumor was analyzed to assess for LOH or a somatic variant in the same gene. **Results:** Of the first 777 clinical cases submitted for concurrent somatic and germline variant analysis, the most common tumor types were prostate (39%), pancreatic (17%), renal cell (12%) and breast (11%) cancers. Since the patients' normal DNA had been sequenced alongside their tumors, we were able to provide rapid germline genetic analysis, with a mean TAT of 6 days. Pathogenic or likely pathogenic germline variants were detected in 22/76 genes, and were found in 157/777 (20.2%) of the cases analyzed. Even after excluding low penetrance alleles such as APC p.Ile1307Lys, 129/777 (16.6%) of cases were found to harbor inherited cancer susceptibility variants. The positive rate varied by cancer type and included 20.4% of prostate, 26.4% of pancreatic, 14.4% of renal cell, and 15.9% of breast cancer cases. Interestingly, we found few BRCA1/2 pathogenic variants in our breast cancer cohort, presumably because high risk patients already had primary genetic testing performed. Conversely, we found BRCA1/2 germline mutations in BRCA-associated cancer types in which primary germline testing might not be frequently ordered, including 11.6% (15/129) of pancreatic cancer, 7.6% (23/304) of prostate cancer, and 33% (3/9) of cholangiocarcinoma cases. Moreover, we genetically diagnosed inherited cancer syndromes, including Lynch syndrome and BAP1 cancer predisposition syndrome, in seemingly sporadic cases. Finally, we demonstrate the biologic importance of these pathogenic germline variants by showing that 48% have a somatic "hit" in their tumor, either an SNV, small indel, or LOH. **Conclusions:** The clinical importance of analyzing the germline for cancer patients cannot be underestimated, as 20% of our patient cohort was found to harbor a pathogenic germline variant for an inherited cancer predisposition syndrome.

**G07. Validation of a 176-Gene Next-Generation Sequencing Panel for Nuclear Encoded Mitochondrial Genes**

*S. Roellinger, J. Balan, Z. Tu, E.W. Klee, D. Oglesbee, L. Hasadsri, W.E. Highsmith*
*Mayo Clinic, Rochester, MN.*

**Introduction:** The majority of mitochondrial proteins (>98%) are encoded by nuclear genes, and mutations in these contribute to multiple mitochondrial disorders. A targeted next-generation sequencing (NGS) panel was validated to sequence 176 clinically relevant, nuclear-encoded mitochondrial genes. **Methods:** Library preparation and target enrichment was performed via a specific, custom probe mix designed to the genes of interest (Agilent SureSelect$^{XT}$). Libraries were sequenced on the Illumina HiSeq using paired end reads. CLC Genomics Server was used for read alignment and variant calling. To demonstrate accuracy, we assessed seven samples with whole exome or whole genome data available and 14 samples with previously reported mutations. In addition, 28 samples with a clinical phenotype consistent with mitochondrial disease but no previously identified mutations were tested. Precision was assessed using three samples run in triplicate both on the same run and across three separate runs. **Results:** We observed an average of 52% reads mapped to target (average of 99.5% reads mapped to genome) and no regions <100X depth of coverage (minimum of ten million paired end reads per sample). Of the 14 samples with previously identified mutations, 100% (22/22) of variants were detected. We observed >98% concordance between our results and variants reported by reference data for samples with whole exome/genome data, excluding homopolymer regions (resolution of discrepant results by Sanger sequencing is ongoing). Deleterious, likely deleterious, or uncertain significance variants were detected in 26 out of 28 samples from symptomatic patients with no previously identified mutations; 11 samples contained ≥1 deleterious/likely deleterious variant and four samples contained ≥2 deleterious/likely deleterious variants. Results were 100% concordant among replicates in the precision study (within and across runs). **Conclusions:** We have demonstrated validity of a 176-gene next-generation sequencing panel to aid in the diagnosis of mitochondrial disorders involving nuclear-encoded mitochondrial genes.

**G08. Computational Confirmation of Maternity and Paternity Using SNP Data from Chromosomal Microarrays**

*D. Xia[1], C. Zhang[2], M.H. Harris[2], Y. Shen[2]*
*[1]Brigham and Women's Hospital, Boston, MA; [2]Children's Hospital Boston, Boston, MA.*

**Introduction:** Germline copy number variants in affected probands detected on array comparative genomic hybridization (aCGH) testing are not infrequently classified as variants of uncertain significance. To further characterize these variants, some laboratories will test the parents of the probands to determine if the variants are *de novo*. For this type of testing, maternity and paternity are often assumed. This is a potentially problematic assumption, however, because the rates of non-paternity in the population may not be negligible. Further, human errors resulting in sample switching can also occur. In this study, we describe the validation of a computer script designed to assess maternity and paternity using single nucleotide polymorphism (SNP) data from Agilent microarrays. **Methods:** The script was written for R (version 3.2.2). Data for 30 thousand SNPs from Agilent microarrays for a proband and his/her putative parent (either mother or father) are combined. The SNPs are then filtered to include only those that are homozygous in both the proband and the parent. For each included SNP, the script determines if the alleles of the proband and parent are identical. If they are identical (*e.g.*, parent is AA; proband is AA), the SNP is concordant. If not identical (*e.g.*, parent is TT; proband is AA), the SNP is discordant. We hypothesized that true proband-parent pairs (true pairs) should have very few discordant SNPs, whereas false proband-parent pairs (false pairs) should have many more discordant SNPs. To test this, we assessed the script using 10 putative trios (10 matched probands, mothers, and fathers) who had aCGH testing at Claritas Genomics (Cambridge, MA). The number of discordant SNPs, number of concordant SNPs, and the proportion of discordant SNPs were tracked across 20 possible true pair and 180 possible false pair combinations. **Results:** The number of discordant SNPs clearly separates the true pairs (median = 39.5, range = 18 to 62) from the false pairs (median = 1129, range = 920 to 2165). The proportion of discordant SNPs also separates true pairs (median = 0.0031, range = 0.0013 to 0.0072) from false pairs (median = 0.092, range = 0.077 to 0.19). On the other hand, although the number of concordant SNPs is statistically different between true pairs and false pairs, there is considerable overlap between these groups. **Conclusions:** Our computer script is able to confirm paternity and maternity using SNP data for the 10 trios tested. Further, by confirming maternity and paternity, pre-analytic sample switching is excluded. Computational SNP analysis can be easily implemented for trio studies performed on Agilent microarrays.

**G09. Clinical Utilization of Genome-Wide Methylation Testing in Pediatric Patients**

*B. Sadikovic[1], K.M. Boycott[2], C. Schwartz[3], G. Pare[4], C. Howlett[5], P. Ainsworth[1], L.C. Schenkel[5]*
*[1]London Health Science Center, London, Ontario, Canada; [2]Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; [3]J.C. Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, SC; [4]McMaster University, Hamilton, Ontario, Canada; [5]Western University, London, Ontario, Canada.*

**Introduction:** The pathophysiology of pediatric Developmental Delay and Intellectual Disabilities (DD/ID) is underscored by genetic and environmental factors, which can both be facilitated by epigenetic mechanisms. DNA methylation is an epigenetic mechanism essential to many biological processes, including transcriptional regulation, establishment of chromatin states, imprinting, X-chromosome inactivation, development and tissue specification, genomic stability and silencing of the repetitive elements. Accordingly, DNA methylation defects have been shown to play a key role in many pediatric and adult onset disorders, including imprinting diseases and carcinogenesis. However, there is little known about genome-wide DNA methylation changes in patients with DD/ID. Our laboratory has performed genome-wide DNA methylation testing in 1000 patients with a wide range of disorders associated with DD/ID to clinically validate the use of a genome-wide DNA methylation assay and define epi-signatures of various DD/ID conditions. **Methods:** DNA methylation of the peripheral blood specimens was assessed using Illumina HumanMethylation450 BeadChip and data was analyzed using a custom algorithm with Partek Genomic Suite software. Methylation patterns of individual patients were compared to normal reference cohort of >300 individuals and were prioritized based on the statistical parameters including methylation difference, *p* value, and F value; and functional parameters including distance to the CpG islands and gene promoter, and the proximity to the known disease-causing genes. **Results:** We have validated this approach for sensitive detection of imprinting disorders including Angelman, Prader-Willi, Beckwith Wiedemann, and Russell-Silver syndrome, as well as Fragile X syndrome. We also discovered novel highly specific diagnostic epigenetic signatures in the peripheral blood of patients with Alpha Thalassemia/Mental Retardation syndrome X-Linked, Floating-Harbor syndrome, Autosomal Dominant Cerebellar Ataxia with Deafness and Narcolepsy (ADCA-DN) and Claes-Jensen X-linked Mental Retardation syndrome. **Conclusion:** These