Platform Presentation

Development of data pipelines for the collection and quality control of clinical variant data.

Michael Cornell[1], Kathryn Robertson[1], and Andrew Devereau[1]

National Genetics Reference Laboratory, Genetic Medicine, 6th Floor, St Mary's Hospital
Oxford Road, Manchester M13 9WL

Curators of clinical variant databases must overcome two large hurdles to develop and maintain high quality resources. The first is obtaining data from the diagnostic laboratories, the second to ensure data quality when data are obtained from a large number of submitters.
Should the curator overcome the problem of persuading laboratories to submit data, they then have to face the issue of those laboratories storing their variants in a multitude of formats e.g. spreadsheets, that must be reformatted and checked by the curator before they can be added to the database, greatly increasing their workload.
The problem of maintaining data quality in public databases such as GenBank, TrEMBL, and KEGG has already been documented. While this can have serious implications, the need to maintain data quality in clinical variant databases is arguably much greater, since this information may directly influence patient diagnosis and treatment. This will become even more important with the development of next generation sequencing based tests as the volumes of data generated will require variant analysis to be automated.

The diagnostic mutation database (DMuDB https://secure.dmudb.net) was established in 2005 to allow the secure sharing of variant data generated by UK diagnostic laboratories. Since September 2011, DMuDB has been available to diagnostic laboratories outside of the UK and is now used by 105 laboratories worldwide. Currently the database contains 45117 variant reports from 78 genes.

As DMuDB has grown the need to streamline submission processes and ensure data quality has become more apparent. ETL (Extract, Transform and Load) workflows have been developed using Kettle (http://kettle.pentaho.com/) to improve data flow into DMuDB. A quality control pipeline using the Taverna Workflow Management System (http://www.taverna.org.uk/) and the Mutalyzer web service (https://mutalyzer.nl/) has been developed to ensure the correct use of HGVS nomenclature. To ensure consistency between submissions, we are conducting regular database audits to identify instances where the pathogenicity of a variant differs between submissions. This has required us to establish protocols for re-contacting users regarding their submissions.

Platform Presentation

Chromosomal Microarray Analysis in a Pediatric Hospital: Ordering Practices and Outcomes

J.A.Kaylor[1], N.Danylchuk[2], J.Vuk[2], T.Lepard-Tassin[2]

[1]Arkansas Children's Hospital; [2]University of Arkansas for Medical Sciences

Background: Chromosomal Microarray Analysis has replaced the karyotype as a first test to detect genetic abnormalities in postnatal evaluations of patients with developmental delay (DD)/intellectual disability (ID), autism spectrum disorders (ASD), or multiple congenital anomalies (MCA), as recommended by the American College of Medical Genetics. Despite this recommendation and previous studies reporting that the 12–13 % diagnostic yield of microarray is superior to a 2–3 % diagnostic yield of karyotype, some health insurance providers exclude coverage of CMA, citing a lack of clinical utility, low diagnostic yield, and the experimental status of the test itself. Though the diagnostic yield of CMA is established for the phenotypes listed above, there is not sufficient data regarding how this test is applied to other phenotypes. Growth delays and speech delays have been recommended by the ACMG as appropriate indications for CMA, but the diagnostic yields have not been widely studied. Other common indications, such as seizures, may be clinically distinct from those phenotypes noted by the ACMG, and there is little data concerning the utility of CMA in these cases. The effectiveness of CMA when ordered by different physician specialties has not been evaluated in detail.
Purpose: The aim of the study was to determine how the CMA is utilized among our facility's physicians and how effectively CMA is used. We hypothesized CMA at the study location would have a diagnostic yield similar to established 12–13 % and that the Geneticists would have a significantly higher yield than other physician specialties.

Methods: We reviewed 393 medical charts in one Mid-South pediatric hospital to establish the ordering practices and outcomes of the CMA when ordered between October 2010 to March 2011 by various physician specialties and for varying indications.

Results: CMA was more frequently ordered for indications other than those recommended by the ACMG than those that are recommended. CMA was ordered for patients with seizures (n=50) more commonly than for patients with autism (n=32) or multiple anomalies (n=32), making it a common indication for testing. CMA was ordered more frequently by other physician specialties (n=272) than by Geneticists (n=121). CMA yielded 15.5% detection of pathogenic variants and 8% detection of variants of uncertain significance. Diagnostic yields were calculated for indications that previous literature has not addressed [23.7 % for Growth Problems (n=23)]. The diagnostic yield for an indication of seizures (12%; n=50) was comparable to DD/ID (13.8%; n=126), ASD (12.5%; n=32) and MCA (12.5%; n=32). Furthermore, the regardless of the specialty of the ordering physician the diagnostic yield remained comparable.

Discussion: These findings provide further evidence of the greater diagnostic yield of the CMA as a first tier genetic test. This was true regardless of the specialty of the ordering physician. The wide use of CMA for different indications and the relatively similar diagnostic yield among indications may provide evidence that CMA could have broader use beyond those indications recommended by the ACMG. The frequency of pathogenic variants or variants of uncertain significance may indicate the importance of proper pre and post-test counseling services.

Platform Presentation

Resources to Enhance the Understanding of Genomic Disorders: Lessons Learned from Patient Registry Reported Data

Vanessa Rangel Miller[1], Kyle Brown[1], Kadi Luchsinger[2]

[1]PatientCrossroads, San Mateo, CA, [2]Dup15q Alliance, Fayetteville, NY

Background: The rarity and diversity of genomic disorders pose unique challenges for research and understanding. Improvements in laboratory technologies increasingly lead to the identification of rare variants for which a clinical prognosis may be limited. This same challenge is seen across the rare disease community, and results in major public health, economic and societal burdens. The ability to combine genetic data with clinical data is more important than ever. To empower patients affected by Dup15q, and many other conditions, PatientCrossroads hosts a registry to collect patient-reported data. By utilizing a curation process, validated registry data has been provided to investigators to facilitate further research. In this abstract, we present the key learnings for patient registry programs to ensure maximum value to the community.

Program: The Dup15q registry uses a web-based HIPAA-compliant PatientCrossroads system, with IRB oversight. Data is curated by genetic counselors to assess the quality of patient accounts, survey responses and submitted genetic test results. Patients create password-protected accounts and consent to share data in a de-identified manner with the greater community. Providing immediate access to the aggregate data and educational content keeps patients engaged.

Results: As of March 2013, 238 people (ages 0-45) registered, 135 male/103 female, from 18 countries. 51% report seizures (over half expected from literature; Dup15q Alliance). 14 of 101 males reported undescended testicles. 18% report ADD/HD and 18% report anxiety. Developmental problems are commonly: 65% report sitting at 8-18 months (10-20 months expected). 68% report walking at 12-30 months (13-54 months expected). 45% have no speech; 27% have a mix of meaningful speech & echolalia/jargon. 55% report autism or ASD. 131 submitted genetic results, of which 74% are idic. Of those, 62% report autism or ASD.

Recommendations: Curated patient-reported data is a valuable source of information to augment the understanding of genomic conditions. We recommend that patient registries for other genomic disorders curate the data, enable patient-controlled consent for data-sharing, return data to patients, and use common data elements to promote pan-disease analysis. To address the challenge of obtaining medical reports and medications, PatientCrossroads has partnered with Informedika to develop a patient-initiated process to electronically transfer medical records to a registry account. To support organizations that cannot afford an independent registry, PatientCrossroads has created the CONNECT program to provide no cost customizable registries. By developing an openly-accessible registry platform, we can change the way patients, researchers and organizations work together.

Platform Presentation

Reporting of incidental copy number variation detected by chromosomal microarray analysis in "normal" parents/family members.

A Patel[1], S Cheung[1], C Bacino[1], P Stankiewicz[1], A Breman[1], S Lalani[1], J Smith[1], C Shaw[1], S Peacock[1], A Braxton[1], Laura Ellis[1], P Ward[1],J Lupski[1], A Beaudet[1], W Bi[1]

[1] Dept of Mol & Hum Genet, Baylor College of Med.

Detection of rare and common copy number variations (CNVs) in the human genome has rapidly increased in the past few years with the widespread adoption of chromosomal microarray analysis (CMA). In an attempt to interpret some of the CNVs identified in probands, CMA on parental or family members is a routine recommendation. Our laboratory has studied approximately 3500 parental samples and in 1% of the cases a clinically relevant CNV was detected that was not seen in the proband. These included CNVs in genes involved in autosomal/X-linked recessive disorders and adult onset disorders such as HNPP (hereditary neuropathy with liability to pressure palsies). Larger CNVs in chromosome regions associated with incomplete/unknown penetrance such as duplications in 1q21, 15q13 and 16p13 were also detected. Mosaic aneuploides for whole chromosomes were detected, including mosaicism for trisomy 3 in one parent. Constitutional mosaic trisomy 3 is very rarely detected postnatally. However, trisomy 3 is found in 20-30% of T-cell non-Hodgkin's lymphoma (NHL) and is also seen in B-cell lymphomas. Additional tests to rule out a hematological malignancy were recommended. CNVs in other adult onset cancer predisposition genes include *PMS2* and *MYCN*. Reporting of these incidental CNVs have a significant impact on risk counseling for future pregnancies and other family members at risk. Currently there are no formal policies for disclosure of incidental findings identified by CMA or Next Generation Sequencing that could inform interventions or improve health outcome. Additionally, who is obligated for the disclosure of incidental finding, the laboratory or the clinician? For consistency a greater regulatory guidance and education of individuals using genetic testing is needed for identifying and reporting of incidental findings.

1. Premature Ovarian Failure: a simple case of Xq deletion or is it more?

Umut Aypar[1], Nicole L. Hoppman[1]

[1]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN

Approximately 40% of premature ovarian failure (POF) is due to genetic factors such as Turner syndrome, Xq deletions, and *FMR1* pre-mutations. Recently, we received a peripheral blood specimen from a 34 year old female for chromosome analysis as well as Fragile X testing due to POF. The initial chromosome analysis revealed an interstitial deletion of Xq from q22.1 to q26, consistent with the patient's POF. In the mean time, the PCR for *FMR1* CGG repeat sizing demonstrated 33 CGG repeats (normal); however, only one copy of the *FMR1* gene was present on the southern blot; inconsistent with the chromosome results since *FMR1* is located at Xq27.3. Therefore, array comparative genomic hybridization (array CGH) was performed to clarify this discrepancy. Array CGH revealed that the patient actually has a terminal Xq deletion, 49.3 megabases in size, from Xq22.3 to Xqter (105,595,496-154,886,101, human genome build 18) as well as a terminal duplication of 20.2 megabases from 4q33 to 4qter (171,058,931-191,250,527, human genome build 18). Metaphase FISH studies using Xq and 4q subtelomere probes demonstrated that these abnormalities are the result of an unbalanced X;4 translocation, which was not obvious by chromosome studies. According to the patient's physician, the patient has normal intelligence and no health problems other than POF. Therefore, it is possible that preferential inactivation of the structurally abnormal X chromosome could result in inactivation of attached chromosome 4 segment, hence minimizing the potential phenotypic effects of the 4q duplication. With respect to family history, the patient has a healthy son; however she has two maternal aunts with intellectual disability. This suggests that her mother may be a carrier of the balanced form of this translocation and the maternal aunts may have the reciprocal unbalanced translocation resulting in 4q deletion and Xq duplication, potentially explaining their intellectual disabilities. It is also possible that they have the same unbalanced translocation as this patient but without favorable X inactivation. This case report highlights the usefulness of array CGH studies in patients beyond those with developmental delay/intellectual disability, autism spectrum disorders and multiple congenital anomalies. Furthermore, array CGH not only clarified the abnormality due to higher resolution compared to chromosome studies but also significantly impacted the recurrence risks for this family.

2. Survey of genetic counselors offering prenatal microarray testing: Opinions and educational needs

Barbara A. Bernhardt,[1] Katherine Kellom,[1] Alexandra Barbarese,[1] W. Andrew Faucett,[2] Ronald J. Wapner[3]

[1]Division of Translational Medicine and Human Genetics, Perelman School of Medicine, University of Pennsylvania; [2] Geisinger Health System; [3]Columbia University, and the NICHD Prenatal Microarray Study Group

Genetic counselors play a key role in the diffusion of prenatal microarray testing by providing pre-test education and counseling, post-test results disclosure, and follow-up counseling. We conducted a survey of members of the National Society of Genetic Counselors who counsel for prenatal microarrays to explore their comfort with prenatal microarray testing, related educational needs, and opinions on returning various types of results. Of the 98 respondents to date, nearly all were comfortable with most aspects of prenatal microarray testing. There was less comfort with some aspects, including counseling about uncertain results (40% uncomfortable), helping a patient decide about pregnancy termination if there is an uncertain result (56% uncomfortable), and obtaining information to clarify the meaning of an uncertain result (34% uncomfortable). All respondents were interested in additional education and training related to prenatal microarray testing. Over 90% wanted education on dealing with the uncertainty of results, communicating abnormal or uncertain results, and locating information about implications of specific CNVs. Respondents also felt it very important that couples continuing pregnancies after receiving positive microarray results receive information and services, including printed materials about the abnormality found (77% very important), access to experts (76%), and a plan for follow-up after the baby is born (86%). Considered less important were ongoing counseling or psychological support (60% very important) and an opportunity to speak with parents of a child with similar findings (47%). Although most respondents agreed it is important to report all possible results to patients (including both de novo and inherited CNVs associated with a 20% risk for autism, and results showing areas of homozygosity consistent with parental consanguinity), there was less support for reporting findings indicating a 20% risk for developing a treatable adult-onset condition (57% agreed) or a 20% risk for an untreatable adult-onset disorder (66% agreed). A slight majority (58%) agreed with returning a result indicating the fetus is a carrier for an autosomal recessive disorder. These survey data document that prenatal genetic counselors would like additional education and resources relevant to prenatal microarray testing for both themselves and their patients, and that there is variability in opinions regarding the return of various types of test results.

Supported by grant 2U01HD055651-06 from the NICHD, National Institutes of Health

3. Reference Resources to Support Clinical and Research Labs

Deanna M. Church[1], Chen Chao[1], Jonathan Trow[1], Maryam Halavi[1], Cliff Clausen[1], Peter Meric[1], Victor Ananiev[1], Hem Vichet[1], Daniel Frishberg[1], Lianghou Wu[1], Victor Joukov[1], Anatoliy Kuznetsov[1], Christopher O'Sullivan[1], Wendy Rubinstein[1], Stephen T. Sherry[1], Madhuri Hegde[2], Birgit Funke[3] and Lisa Kalman[4]

[1]National Center for Biotechnology Information, Bethesda, MD; [2]Department of Human Genetics, Emory University, Atlanta, GA; [3]Harvard Medical School, Boston, MA; [4]Centers for Disease Control and Prevention, Atlanta, GA

Advances in Next Generation Sequencing (NGS) technology are shifting how we think of variant detection and molecular diagnostics. The high throughput and low cost of NGS enables analysis of an entire exome or genome. This approach is not without difficulties. Each NGS platform has different error profiles and there are a myriad of alignment and analysis tools available for use. To fully realize the promise of NGS for variant analysis, in both the research and clinical lab, the need for robust reference data is critical.

To this end we have developed reference resources for two HapMap samples, NA12878 and NA19240. As part of the Genetic Testing Reference Materials Coordination Program (GeT-RM) a number of clinical, commercial and research labs have performed their NGS-based tests on one or both of these samples. These tests included targeted gene tests, whole exome tests and whole genome tests. In all cases, labs reported their variant call data, and in many cases we have additional data including NGS sequence reads and alignments, Sanger sequence data, validation information and capture reagent information. We have combined this wealth of data into a single resource allowing for bulk data download as well as for a genome browser. This allows users to compare results obtained in their own lab with results of the GeT-RM project and will facilitate assay development, validation, and quality control. We will discuss some of the challenges of combining data from different laboratories to produce a centralized resource.
Additionally, we wish to support computation development in the area of sequence alignment and variant calling. To this end we have created an instance of the Amazon Elastic Compute Cloud (EC2) containing all publicly available sequence data for the NA12878 sample. This resource will provide a fertile development environment for software developers as they look to improve analysis tools and pipelines.

This work was supported in part by the Intramural Research Program of the NIH, NLM

4. Copy Number Profiling of Single Cells Using Oligo aCGH

Paula Costa[1], Stephanie Fulmer-Smentek[1], Peter Webb[1], Emily Leproust[1], Anniek De Witte[1]

[1]Agilent Technologies Genomics, Santa Clara

Genomic imbalances are a major cause of constitutional and acquired disorders. Due to genomic heterogeneity and sample availability, the characterization of individual genomes is becoming an area of focus in cancer, reproductive and stem cell research. FISH and PCR-based techniques have traditionally been used to study single cells, but their application is limited by the number of loci that can be analyzed simultaneously. More recently, BAC arrays, which suffer from low resolution and poor reproducibility, have also been applied to single cell analysis. Taking advantage of the high-resolution and sensitivity of the Agilent SurePrint G3 Microarrays technology, we applied the power of the array Comparative Genomic Hybridization (aCGH) platform to copy number (CN) profiling in single cells.

Agilent's aCGH application relies on a two-color assay to measure CN changes in a test sample relative to a reference. For proof-of-principle experiments, genomic DNA (gDNA) from a sample with a known amplification on the short arm of chromosome 9 was diluted to single cell levels. To assess CN changes in true individual genomes, single cells were biopsied from embryos. Whole genome amplification (WGA) was performed in all test and reference samples to increase the amount of DNA while maintaining the genomic representation. Test and reference amplified DNA samples were differentially labeled with Cy5 and Cy3 dyes, combined and hybridized for 16 hours to SurePrint G3 Human Catalog 8x60K CGH Microarrays, with uniform backbone coverage and denser coverage in genes. Following hybridization, the images were scanned, the data extracted and analyzed for CN alterations using algorithms implemented in Agilent CytoGenomics.

With an optimized workflow for sample and microarray processing, along with an optimized reference, the expected amplification on the short arm of chromosome 9 was identified in the gDNA sample diluted to single cell levels and subjected to WGA. CN aberrations, both whole chromosome losses or gains and smaller aberrations of portions of chromosome arms, were assessed in single cells biopsied from embryos. A reduction in the noise level and improvement in the data was observed for samples hybridized to a pooled reference as compared to a single reference. Using an approximately 24-hour single cell workflow from sample preparation to analysis, reliable results can be obtained at a high resolution while remaining cost-effective.

5. Capturing phenotypes in clinical variant databases

Maja Milicic-Brandt[1], Andrew Devereau[1], Kathryn Robertson[1], Michael Cornell[1]

[1]National Genetics Reference Laboratory; Genetic Medicine, 6th Floor, St Mary's Hospital
Oxford Road, Manchester M13 9WL

Next generation sequencing (NGS) is bringing new challenges on how to store and interpret massive quantities of genetic data. Storing patient phenotype data in clinical variant databases in a standard form will be essential for: (i) discovery of genotype-phenotype correlations; (ii) sharing and merging of phenotype data from different sources; (iii) differentiation of new syndromes; (iv) finding frequency of phenotypes with certain disorders; (v) prediction of disease severity based on phenotype.

The Human Phenotype Ontology (HPO http://www.human-phenotype-ontology.org/) contains over 10000 terms describing phenotypic abnormalities which are related to each other by subclass relations (e.g. based on anatomy). Since HPO is the most comprehensive, best structured and best maintained terminology of human phenotypes, it has emerged as a de-facto standard in this area.

We are working on adding patient phenotype data based on HPO to our Diagnostic Mutation Database (DMuDB https://secure.dmudb.net). This project relies on two separate sub-projects: (1) phenotype data capture by sophisticated electronic clinical referral forms and (2) extension of the database in order to store phenotype data.

There is an extremely wide range of genetic diseases and specialists from different areas have diverse methods of collecting phenotype data. In the area of developmental disorders, clinicians prefer not to have any "obligatory" fields in the forms, but to choose from the full set of HPO terms which is well suited for encoding relevant abnormalities. On the other hand, in many other areas, clinicians are used to forms asking for specialised data relevant for their field. A large part of this data is either quantitative or too fine-grained to be properly captured by HPO. Such data is out of scope of DMuDB and needs to be approximated by the "closest" HPO terms.

When storing HPO terms in our variant database we need to allow for sequences of HPO codes representing a phenotypic abnormality and a set of qualifiers such as age of onset or speed of progression. For example, we can describe early-onset breast carcinoma by the HPO terms "Breast carcinoma" and "Young adult onset". Many HPO terms such as "Adult onset sensorineural hearing impairment" already contain qualifying information and such terms are more convenient for data exchange with other systems. Finally, when it comes to answering database queries, we plan to exploit HPO phenotype taxonomy in order to get better answers.

6. A case study of *GRIK1* deletion in a patient with seizures

Denae M. Golden[1], Dani L. Bishay[1], Diane L. Pickering[1], Bhavana J. Dave[1], Warren G. Sanger[1]

[1]Human Genetics Laboratories, University of Nebraska Medical Center and the Munroe-Meyer Institute for Genetics and Rehabilitation, Omaha, Nebraska, USA

The *GRIK1* gene, located at 21q21.3, is activated during normal neurophysiologic processes in the mammalian brain. It is an ionotropic glutamate receptor which is a known excitatory neurotransmitter receptor. Variation within glutamate receptors has long been associated with seizure activity; however, there is a paucity of literature and research related to *GRIK1*, with only one report suggesting a possible association between mutations within the *GRIK1* gene and juvenile absence epilepsy (JAE). The findings supported an allelic association between variation of the *GRIK1* tetranucleotide repeat and JAE.

We present a case of an eight year old male with prenatal polysubstance exposure, psychiatric concerns, absence seizures and two episodes of generalized seizures with a normal EEG. Genetic testing included Fragile X testing (performed by another lab), high resolution chromosome studies, and Postnatal 180K Exon Microarray studies. The Fragile X and the chromosome studies were normal, however, the microarray study revealed a 94Kb deletion partially covering the 5' terminal end of the *GRIK1* gene at 21q21.3. The linear location of the deletion was 31,270,211bp – 31,364,532bp (hg19). An aberration of this gene region had not been observed previously by our laboratory. Very little overlap in the Database of Genomic Variants was noted, and no other cases of this size were reported in dbVar.

Given the clinical correlation of this patient's phenotype of seizures, the single previously reported study, and the position of the deletion (5' end of *GRIK1*) it is likely that it includes the tetranucleotide repeat portion of the gene which is associated with JAE. Currently there is no mention of phenotypes associated with deletions of *GRIK1*. Unfortunately parental studies are not possible because this patient is adopted and the biological parents are unavailable. Since this gene region is understudied, there is a lack of informative literature to help guide us in our interpretation of this aberration. However, since mutation in this gene has been linked to susceptibility to JAE and there was a clinical correlation with seizures in this patient, our results of deletion within the *GRIK1* were interpreted as "uncertain but likely pathogenic". Further studies to help understand the clinical similarity of the outcome of a mutation versus a deletion within this gene will better refine the interpretation of the results.

7. Detection of Rare Familial 17p13.3 Microduplication in a Patient with Developmental Delay and Behavioral problems

M. Hankerd[1], M. Kristofice[1] , R. Chikarmane[3], D. Stockton[3] , S. Ebrahim [1, 2].

[1]Detroit Medical Center University Laboratories, Cytogenetics; [2]Department of Pathology, Wayne State University School of Medicine, Detroit, MI; [3]Division of Genetic and Metabolic Disorders, Children's Hospital of Michigan, Detroit, MI.

Variable sized duplications of the 17p13.3 region encompassing the YWHAE and surrounding genes have been reported in few patients with autistic features, behavioral problems, speech delay, motor delay, obesity, subtle dysmorphic features and some hand and foot malformations. We report on the chromosomal microarray evaluation of a 15 year old boy presenting with short stature, obesity, developmental delay, behavioral problems, and a history of osteochondroma. The patient has a 12 year old bother with ADD whom is also being considered for a diagnosis of Asperger syndrome as well as microarray evaluation. A three generation family history was notable for osteochondroma and mental health issues on the mother's side but was otherwise unremarkable. Chromosomal microarray analysis detected a 53.65 kb duplication/copy number gain of chromosome 17p13.3 short arm region involving one OMIM gene (YWHAE). Maternal chromosomal microarray was performed and the mother was found to carry the same 17p13.3/YWHAE duplication. Therefore, the duplication identified in this patient is of maternal origin. No additional family studies have been performed at this time. Capra et al (2012) reported a family with developmental delay and behavioral problems, segregating a 17p13.3 duplication (about 329.5 kb) involving the YWHAE gene and two additional genes. The duplication observed in our patient, however, contains only the YWHAE gene and is smaller than previously described 17p13.3 duplications. Although the duplication detected in our patient is smaller, the clinical phenotype is similar to that described in larger duplications of the 17p13.3 region. This suggests that the YWHAE gene may directly contribute to the abnormal phenotype observed in our patient as well as that seen in the larger duplications of this region. To our knowledge this the first reported case involving only the YWHAE gene specifically. Additional cases are needed to further establish the role of the YWHAE gene in 17p13.3 microduplication syndrome.

8. Are Alterations in the *CHL1* Gene Associated with a Neurodevelopmental Phenotype?

Hashimoto S[1], Erdman L[1], McKinney A[1], Ramsey S[1], Weber C[1], Weslow-Schmidt J[1],Lamb Thrush D[1,2], Atkin J[2,3], Baum R[2], Crowe C[4], Hickey SE[2,3], McBride KL[2,5], Miller M [6], Pfau RB[1], Astbury C[1,7], Gastier-Foster JM[1,2,7], Reshmi S[1,5], Pyatt RE[1,7]

[1]Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, Columbus, OH; [2]Department of Pediatrics, The Ohio State University, Columbus, OH; [3]Division of Molecular and Human Genetics, Nationwide Children's Hospital, Columbus, OH; [4] Dept of Pediatrics, Metrohealth Medical Center, Case Western Reserve University, Cleveland, OH; [5] Center for Cardiovascular and Pulmonary Research, Nationwide Children's Hospital, Columbus, OH; [6] Dayton Children's Medical Center, Dayton, OH; [7]Department of Pathology, The Ohio State University, Columbus, OH

Terminal deletions of chromosome 3 range in size from one to several megabases and are associated with developmental delay, congenital heart defects, dysmorphic features, microcephaly, and low birth weight. The *CHL1* gene, located at 3p26.3, and has been proposed as a candidate gene contributing to the neurodevelopment phenotype associated with this these deletions due to its high levels of expression in the brain. Additionally, there is one report describing an individual with intellectual disability and a balanced translocation [46,Y,t(X;3)(p22.1;p26.3)] with a breakpoint within intron five of the *CHL1* gene (Frints et al. 2003). While there are some reports of copy number variation within the *CHL1* gene region in the Database of Genomic Variation, clinical reports of alterations less than one megabase in size are rare. A pair of siblings has been reported with features including microcephaly, learning difficulties, and intellectual disability who inherited a deletion involving the *CHL1* gene from their asymptomatic father (Cuoco et al. 2011). The ISCA database contains 12 entries (5 gains, 7 losses) for alterations of less than one megabase in size which contain only the *CHL1* gene. Ten of these alterations were reported as being of uncertain clinical significance, one was reported as benign, and one as pathogenic. We describe the molecular and clinical characterization of 6 patients with alterations in the *CHL1* gene identified by clinical oligonucleotide array comparative genomic hybridization (array-CGH) and consider the role of *CHL1* alterations in neurodevelopmental phenotypes. Probands in this series were referred for indications including microcephaly, growth retardation, developmental delay, congenital anomalies, and agenesis of the corpus callosum. Array-CGH identified partial (n=4) and complete gene duplications (n=1); (ranging from 197kb to 330kb), and an intronic deletion (n=1; 54kb). Three of these findings were determined to be paternally inherited (two gains and one loss) and one maternally inherited (one gain) from reportedly normal individuals. A similar intronic loss and a duplication of exons 16 through 28 were identified as incidental findings in two reportedly normal parents (one father, one mother) of probands lacking *CHL1* alterations. While no consistent neurodevelopmental phenotype was evident, the analysis of a larger cohort may be warranted to refine the clinical features of *CHL1* gene alterations.

9. dbVar: A Public Archive for Genomic Structural Variation

Timothy W. Hefferon[1], John Lopez, John Garner, Chao Chen, George Zhou, and Deanna M. Church

National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD

Genomic structural variation plays a central role in genetic variation as it relates to human development and disease. Due to its complex nature, the precise definition of structural variants can be challenging. Centralized, stable public archives like dbVar (and its sister database DGVa, at EBI) are of fundamental importance to enabling investigators to make sense of an ever-expanding stockpile of inter-related genetic variation and clinical data.

Currently dbVar holds structural variation from 99 studies (mostly human, plus 10 additional species). Among our more recently accessioned datasets are two landmark studies: the 1000 Genomes Project (Phase I), and the Sanger Institute's Catalogue of Somatic Mutations in Cancer (COSMIC). Over the past 12 months dbVar has implemented a number of significant improvements which increase both its usability and its value as an archive. We conducted an overhaul of our submission template and database schema, making it easier for users to submit while ensuring we get the data necessary to allow users to get maximal benefit. Another notable improvement is the implementation of a dbVar-specific genome browser which has facilitated the comparison of data between any desired subset of studies; it will also serve as the basis for development of a NCBI-wide variation browser. Lastly, we have greatly improved our ability to capture and display very complex variation data, as exemplified by the sorts of large, multi-step rearrangements that are frequently observed in cancer. More information about dbVar can be found on our website (http://www.ncbi.nlm.nih.gov/dbvar/) and in a recent article (Lappalainen I, et al. (2013) Nucleic Acids Res. Jan;41:D936-41. PMID:23193291).

10. Optimising targeted NGS panels for clinical translational research

Simon Hughes[1], Jennie Bell[2], Chris Mattocks[3], Spencer Howell[1], Jolyon Holdstock[1], Pauline Rehal[2], Matt Lyon[3], Matt Smith[2], Brendan O'Sullivan[2], Michael Griffiths[2], Nick Cross[3], James Clough[1]

[1]Oxford Gene Technology, Oxford, UK. [2]West Midlands Regional Genetics Laboratory, Birmingham, UK. [3]Wessex Regional Genetics Laboratory, Salisbury, UK. Simon.Hughes@ogt.com

Targeted re-sequencing of small panels of genes offers many benefits for clinical research — it is cost effective, allows detection of all variants associated with many inherited diseases, enables ultra-deep sequencing for detection of low frequency variants in heterogeneous samples and reduces the number of unsolicited findings which may be unwelcome. It is important, however, to optimise the initial enrichment step, to avoid introducing bias or error into the sequencing data. We have designed and optimised a targeted, 58-gene panel for molecular profiling of solid tumours. We evaluated four enrichment technologies to see which performed best in a technical comparison: uniformity of sequence coverage, concordance between fresh and FFPE, amount of starting material required, ease and speed of protocol. The top two then went into initial validation with a series of FFPE samples with known but blinded genotype and varying proportions of tumour present.

Agilent SureSelect™ (with Oxford Gene Technology optimised bait design) and Agilent HaloPlex™ gave the best results in the first round. Optimised bait design for SureSelect significantly increased the uniformity of sequence coverage (>90% of bases covered at 0.2X mean). In general, SureSelect performed better than HaloPlex, offering excellent concordance between fresh and FFPE using starting amounts of 500ng DNA.

Crucially, this SureSelect-based Solid Tumour Panel successfully detected all variants present in the blinded series. This included two variants present in KIT, and three in KRAS. Two of the samples with KRAS mutations had small proportions of tumour present (20% and 5% respectively), but the assay still confidently detected both variants. The HaloPlex-based assay only detected 4 of 6 variants, missing both variants which were present at low frequency.

We present this data and discuss the importance of design and choice of enrichment assay format for clinical and translational research NGS applications. The knowledge and experience gained in developing and optimising this capture panel will be used for the development of research panel products in both the constitutional and oncology areas in the future

This research project is partly funded by the Technology Strategy Board, U.K. under the Stratified Medicines for Cancer Programme. The Solid Tumour Panel is intended for Research Use Only. Not for use in Diagnostic Procedures.

11. Intragenic deletion of *STK32B* in a family with isolated cleft palate

Kantarci S, Bui P, Li WL, Zaman NH, Dipple KM

Cleft palates (CP) with a multifactorial etiology involving both genetic and environmental factors are among the most common congenital defects. The size and position of the cleft varies and it may involve only the hard palate, only the soft palate, or both the hard and soft palate. The overall incidence of isolated CP with no other birth defects is about 6 per 10,000 live births in the U.S. Several studies showed genomic copy number alterations, either gain or loss, are involved in the occurrence of cleft palate.

In this study, we report a 2.5 year-old male patient with a cleft of the soft palate and mild mandibular hypoplasia. He underwent a cleft palate repair at age 1.5 years. The patient is developmentally normal and has some misarticulation, but no significant hyponasal or hypernasal speech. Since 9 months of age he has also had frequent infections of otitis media and has ear tubes. His mother also had a cleft palate repair and still has a hypernasal speech.

Chromosomal SNP microarray analysis (CMA) using Affymetrix SNP Array 6.0 platform in the patient's DNA sample revealed a 185 Kb heterozygous deletion of chromosome 4p16.2: arr[hg19] 4p16.2(5,134,473-5,319,774)x1. The overlap with a rare copy number loss variation is ~68% (variation_8995; 1/776 controls=0.13%). The parental CMA testing using CytoScan HD Array platform showed that the mother has a similar deletion: arr[hg19] 4p16.2(5,133,065-5,319,773)x1.

This deletion interval encompasses exons 2 and 3 of *STK32B* (serine/threonine kinase 32B), a gene of unknown function. The *STK332B* gene with 12 exons encodes for a protein with 414 amino acids, which is a member of the serine/threonine kinase family. It is predicted to contain a protein kinase domain (23-283 aa) and two ATP binding domains (29-37 aa and 52 aa). The deletion of exons 2-3 removes the predicted ATP binding domain and thus, it might disrupt the potential kinase function of the protein. A fine mapping SNP linkage study in case-parent trios from four populations for both isolated CL/P (cleft lip with or without cleft palate) and CP on chromosome 4p16 revealed an evidence for linkage in the presence of disequilibrium in several genes, including *STK32B*.

Here we report the first intragenic deletion of *STK32B* in the pathogenesis of isolated cleft palate in an affected family. This study not only supports the association of *STK32B* with CP but also highlights the significant role of CMA in genetic evaluation of CP.

12. Reviewing Laboratory Send Out Genetic Tests: The financial impact of a Genetic Counselor in a pediatric hospital

C.Sailey[1,2], J.A.Kaylor[1],
[1] Arkansas Children's Hospital; [2] University of Arkansas for Medical Sciences

Background: While a majority of Genetic Counselors remain in a clinical, patient care setting, there is an increase in those who serve a non-clinical role. This growing sect, now 23%, defines their primary role including laboratory services and genetic testing. It is known that a Genetic Counselor in the role of reviewing genetic testing for appropriateness has fiscal advantages, though this was proven in a reference laboratory setting (1).

Genetic testing is also increasing on average 15-20% each year in some institutions (2). Our institution's reference laboratory costs (mostly genetic) have nearly doubled from $3 Million to $5.8 Million in the fiscal years of 2008 and 2012, respectively. In the 2011 fiscal year at our institution, the clinical laboratory sent 3,737 genetic tests ranging from $84 to $11,320 to 26 different reference laboratories.

The impact of a Genetic Counselor serving that role in a pediatric institution has been demonstrated in one study, where a Genetic Counselor and a team of PhDs were able to reduce inappropriate genetic testing with savings of $135, 696 in one year through methods of intervention. The authors in that study reviewed only tests that met specific financial or clinical criteria (2). Tests meeting those criteria constituted a portion of all send out testing. There have not been reports of the financial impact when all send out genetic testing is subject to review. Additional reports are needed to provide evidence of a Genetic Counselor's effectiveness in this role. Other aspects of test coordination may be provided by Genetic Counselors, such as consolidating testing to fiscally appropriate reference laboratories.

In addition to the financial impact of a Genetic Counselor in the pediatric laboratory setting, there are opportunities to enhance appropriateness of testing strategies and to supplement reference laboratories and databases with clinical information and the correlation for diagnostic study. In 2012, there was a call to action for Genetic Counselors to serve as agents for providing phenotypic data that aids in test interpretation (3). This role could be filled by a Genetic Counselor reviewing genetic test orders as phenotypic analysis is often a component of the review. By supplying additional clinical information to reference laboratories and databases, the interpretation of testing is enhanced. A Genetic Counselor in the role of reviewing genetic testing may also impact clinical care by educating non-genetics professionals on the significance of certain results, beyond the reference laboratory reports.

Purpose: The aim of our project is to analyze the fiscal implications of a Genetic Counselor's involvement in reviewing genetic testing at a pediatric hospital.

Methods: Genetic testing orders were documented from July 2012 to present. A total of 1448 genetic tests were reviewed by Laboratory Genetic Counselor (number will grow by time of presentation). Medical charts were reviewed for all tests ordered, and phenotypic data were used to determine the appropriateness of a test order. Tests were deemed appropriate or inappropriate and contact with the ordering physician was made and documented to discuss testing. Intervention methods were used to initiate contact and facilitate discussion. Tests were only changed or cancelled with ordering physician approval. Test modifications and cancellations were recorded when applicable.

Results: Of the 1448 tests reviewed as of February 20, 2013, 3% (n=43) were changed from original order to more appropriate test or laboratory. An additional 4% (n=62) tests were cancelled. The fiscal impact of cancellations and changes amounted to $135,067 (number will grow by time of presentation). This number is

comparable to previous reports (2). This reduced the cost of send out testing by 4% based on the fiscal year 2012 estimates. Reasons for cancellations included inappropriate test for patient phenotype, incorrect entry of test into medical system, and incorrect interpretation of the physician's orders.
While genetic tests were reviewed, the choice of reference laboratories became a subject of review. Specifically, analysis of reference labs for commonly ordered tests resulted in a vendor change saving $45,350 yearly, an estimate based on previous yearly orders.

Discussion: A Genetic Counselor is effective in the role of reviewing genetic testing orders. This is demonstrated by the fiscal impact of $135,067 reduction from test alterations and $45,350 reduction from reference laboratory consolidation. These reviews resulted in a total 5% reduction of the send out testing cost for the same time period. Review process identifies more than inappropriate testing. Compared to a previous study in which only select tests were reviewed, the increased financial impact of this study indicates that when possible, a review of all tests results in larger savings. Prior to Genetic Counselor involvement, opportunities such as reference laboratory consolidation and the provision of clinical information to reference laboratories went unexplored.

References:
1. Christine E. Miller, Patti Krautscheid, Erin E. Baldwin, Danielle LaGrave, Amanda Openshaw,; Kim Hart, and Tanya Tvrdik. (2011) Value of Genetic Counselors in the Laboratory.
2. J.H.Conta, B.L.Cole, J.Dickerson, R.M.Jack, M.Wellner, M.L.Astion (2012) Presented A Strategy to Improve the Quality and Economics of Laboratory Send-outs at Seattle Childrens Hospital: The Role of the Laboratory Genetic Counselor. Abstracts from the Thirty First Annual Education Conference of the National Society of Genetic Counselors (Boston, MA, October 2012) Journal of Genetic Counseling December 2012, Volume 21, Issue 6, pp 884-984
3. Karen E. Wain, Erin Riggs, Karen Hanson, Melissa Savage, Darlene Riethmaier, Andrea Muirhead, Elyse Mitchell, Bethanny Smith Packard, W. Andrew Faucett (2012) The Laboratory-Clinician Team: A Professional Call to Action to Improve Communication and Collaboration for Optimal Patient Care in Chromosomal Microarray Testing. Journal of Genetic Counseling October 2012, Volume 21, Issue 5, pp 631-637

13. AutDB: A modular database for accelerating ASD genetic research

Usha Kuppuswamy[1], Eric Larsen[1], Sharmila Banerjee-Basu[1]

[1]MindSpec, McLean, Virginia, United States of America

A major focus of research in the post-genomic era is to decipher the heterogeneous genetic landscape underlying the pathogenesis of complex human diseases such as Autism Spectrum Disorders (ASD). A number of genes in which rare and/or common genetic variants thought to potentially play a role in ASD onset and pathogenesis have been identified. The advent of new techniques such as next generation sequencing (NGS) has resulted in a significant increase in the number of ASD genes with rare genetic variants. With the accelerated growth of genetic data obtained from ASD individuals adding to the already complex genetic landscape of this disease, there is a critical need for databases specialized in the storage and assessment of this data. The autism genetic database AutDB (http://autism.mindspec.org/autdb/Welcome.do) was developed to serve as a publically available web-based modular database for the on-going curation and visualization of ASD candidate genes. Since its release in 2007, AutDB has been become widely used by individual laboratories in the ASD research community, as well as by consortiums such as the Simons Foundation, which licenses it as SFARI Gene. AutDB has been designed using a systems biology approach, integrating genetic information within the original Human Gene module to corresponding data in subsequent Animal Model, Protein Interaction (PIN) and Copy Number Variant (CNV) modules. The number of ASD susceptibility genes in the Human Gene Module of AutDB has increased from 304 genes in December 2011 to 528 genes in March 2013, which demonstrates both the continued discovery of ASD candidate genes and the ongoing curation of these genes into AutDB. In addition, the usage of NGS techniques has contributed to a dramatic increase in the number of rare variants identified in ASD candidate genes (from 1202 in Dec 2011 to 2799 in March 2013) compared to common variants (from 534 to764 over the same period). Functional profiling of ASD genes with rare variants provides insight into the enriched molecular functions of these genes, including ionotropic glutamate receptor binding, serotonin receptor activity, beta-tubulin binding, and voltage-gated sodium channel activity. AutDB serves as a valuable resource for understanding the ever-evolving genetic landscape of ASD and provides researchers with information useful in bioinformatics analyses such as those described above that will aid in unraveling the molecular mechanisms underlying the disease.

14. Prioritization of copy number variation loci associated with autism from AutDB – an integrative multi-study genetic database

Eric C. Larsen[1], Idan Menashe[1, 2], and Sharmila Banerjee-Basu[1]

[1]MindSpec, 8280 Greensboro Dr, Suite 150, McLean, VA 22102, USA; [2] Department of Public Health, Faculty of Health Sciences, Ben Gurion University of the Negev, Israel

Copy number variants (CNVs) are thought to play an important role in the predisposition to autism spectrum disorders (ASD). However, their relatively low frequency and widespread genomic distribution complicates their accurate characterization and utilization for clinical genetics purposes. Here we present a comprehensive analysis of multi-study, genome-wide CNV data from AutDB (http://mindspec.org/autdb.html), a genetic database that accommodates detailed annotations of published scientific reports of CNVs identified in ASD individuals. Overall, we evaluated 4,926 CNVs in 2,373 ASD subjects from 48 scientific reports, encompassing ~2.12×10^9 bp of genomic data. Remarkable variation was seen in CNV size, with duplications being significantly larger than deletions, ($P = 3 \times 10^{-105}$; Wilcoxon rank sum test). Examination of the CNV burden across the genome revealed 11 loci with a significant excess of CNVs among ASD subjects ($P < 7 \times 10^{-7}$). Altogether, these loci covered 15,610 kb of the genome and contained 166 genes. Remarkable variation was seen both in locus size (20 - 4950 kb), and gene content, with seven multigenic (≥3 genes) and four monogenic loci. CNV data from control populations was used to further refine the boundaries of these ASD susceptibility loci. Interestingly, our analysis indicates that 15q11.2-13.3, a genomic region prone to chromosomal rearrangements of various sizes, contains three distinct ASD susceptibility CNV loci that vary in their genomic boundaries, CNV types, inheritance patterns, and overlap with CNVs from control populations. In summary, our analysis of AutDB CNV data provides valuable insights into the genomic characteristics of ASD susceptibility CNV loci and could therefore be utilized in various clinical settings and facilitate future genetic research of this disorder.

15. Chromosome 19p13.3 deletion, including the Peutz-Jeghers locus (*STK11*), in a three-month-old female with congenital heart defects and abnormal brain MRI: A case of 19p13.3 contiguous gene deletion syndrome

Joel A. Lefferts[1], Laura J. Tafe[1], Stephanie E. Vallee[2], Heather B. Steinmetz[1], Elizabeth I. Reader[1], Gregory J. Tsongalis[1], Mary Beth P. Dinulos[2]

Dartmouth-Hitchcock Medical Center and Geisel School of Medicine at Dartmouth, Department of Pathology[1] and Department of Pediatrics, Section of Medical Genetics[2], Lebanon, NH.

Introduction: Peutz-Jeghers syndrome (PJS) is an autosomal dominant condition caused by mutations in the serine/threonine kinase 11 gene (*STK11;* located within 19p13.3) and is characterized by gastrointestinal polyposis, mucocutaneous pigmentation and a predisposition for various types of cancer. In PJS the *STK11* gene can be disrupted by single nucleotide variants or deletions involving all or part of the gene. Recent reports describe larger contiguous gene deletions of approximately 1.1 Mb within 19p13.3 including *STK11* and approximately 50 other genes in patients presenting with phenotypes that expand beyond classical features of PJS. These additional phenotypic characteristics include borderline intellectual disabilities, hypotonia, seizures, mildly dysmorphic facial features, weight and height at or below 10th percentile, as well as variably expressed cleft palate and congenital heart defects. We report a smaller, overlapping deletion in a patient with a similar clinical presentation.

Methods: Chromosome microarray analysis was performed using the CytoScan® HD Array (Affymetrix, Santa Clara, CA) with DNA extracted from peripheral blood of a three-month old female. The indications for testing were abnormal brain MRI and congenital heart defects (ASD/VSD).

Results: Chromosome microarray analysis detected a 750 kb single copy loss of approximately 35 genes, including *STK11*, within cytoband 19p13.3. A subsequent genetic evaluation was significant for findings of hypotonia, developmental delay, and low weight for her length, in addition to the congenital heart defects and brain MRI abnormalities. Parental testing was recommended and is currently pending. Following recommended surveillance guidelines for PJS, this patient will be followed over time for the development of gastrointestinal polyps and other features of PJS.

Conclusions: The findings in our patient are similar both in clinical presentation and gene content to recent reports of patients with expanded Peutz-Jeghers phenotypes. While the lack of typical Peutz-Jeghers features in our patient is expected due to her young age, they are predicted to develop during childhood. The smaller size and decreased gene content of the deleted region in our patient may provide insight into the critical region and gene(s) responsible for this emerging 19p13.3 contiguous gene deletion syndrome that includes the *STK11* gene and an expanded Peutz-Jeghers phenotype.

16. ClinVar, GTR, MedGen and more: how NCBI supports representation of information about medically important variation

Donna Maglott[1], Deanna Church[1], Jennifer Lee[1], Melissa Landrum[1], George Riley[1], Doug Hoffman[1], Brad Holmes[1], Stephen Sherry[1], Wendy S. Rubinstein[1]

[1]National Center for Biotechnology Information, NIH, Bethesda, MD 20892

During the past year, NCBI has brought many resources on line to support management of information about medically important variation and its relationship to phenotype. These include ClinVar (www.ncbi.nlm.nih.gov/clinvar/), the NIH Genetic Testing Registry (GTR, www.ncbi.nlm.nih.gov/gtr/) and MedGen (www.ncbi.nlm.nih.gov/medgen/). These interconnected databases and their associated services provide standardized representation of phenotype (MedGen); archive and version records of the interpretation of rare variation in a clinical context (ClinVar); and accession genetic tests, report quality measures for those tests (*e.g.* analytical validity, clinical utility and proficiency testing), and facilitate access to practice guidelines (GTR). These resources are also tightly coupled with other familiar resources at NCBI, including dbSNP, dbVar, Gene, PubMed, and the sequence databases.

We will provide a brief overview of the development of these resources, namely consulting with stakeholders, refining the data model, processing submissions, and developing the web interfaces. We will concentrate on ClinVar as an example of leveraging the rich resources at NCBI to package information for the medical genetics user community. We will show how ClinVar's model of applying standard representations of phenotype, genotype, interpretation, and evidence facilitates evaluation of the interpretation of rare variation from multiple submitters. We will also review the data model, data flows, current contents, how to access the data, tools to analyze data locally, lessons learned, and unresolved issues.

17. The ISCA Project at NCBI

Jose Mena, Erin Riggs, Erin Kaminsky, Jack Wang, Natalia Popova, John Lopez, Tim Hefferon, John Garner, Cliff Clausen, Victor Joukov, Anatoliy Kuznetsov, Mike Feolo, Christa Lese Martin, Deanna M. Church

The International Standards for Cytogenomic Arrays (ISCA) project provides a centralized online repository for genotype and sample-level phenotype data related to patients with developmental delay, autism spectrum disorders, and/or congenital anomalies. Clinical labs performing cytogenomic array testing for a variety of indications submit data to the database of Genotypes and Phenotypes (dbGaP). To assist submitters in their efforts to accurately annotate each CNV with a clinical interpretation, a conflict checking and resolution tool has been developed that ensures all overlapping CNVs from their lab are annotated in a similar way. Conflicts with a gold standard dataset and inconsistencies among a submitting lab's annotations are presented in separate reports in a web browser; the web interface allows for convenient and quick conflict resolution.

The CNV calls are de-identified and uploaded to a public browser in the database of genomic structural variation (dbVar), where they are organized into three kinds. The first is a gold-standard list of CNVs drawn from literature review whose clinical significance is well established (http://www.ncbi.nlm.nih.gov/dbvar/studies/nstd45/). The second is a broad collection from many laboratories of CNVs from patients presenting with developmental delay who request genetic testing (http://www.ncbi.nlm.nih.gov/dbvar/studies/nstd37/). The final is a collection of CNVs detected in prenatal samples whose mothers present with a clinical indication associated with developmental delay (http://www.ncbi.nlm.nih.gov/dbvar/studies/nstd75/). Individual-level data related to these resources are kept in dbGaP (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000205.v5.p2).

The ISCA project is comprised of 20,291 CNVs from 22,037 patients for a total of ~2.75 billion genotypes. Summary data are publicly available in the database of genomic structural variation (dbVar) and comprehensive data is available to authorized users in the database of Genotypes and Phenotypes (dbGaP). Our suite of resources has proven very useful for both clinical testing labs and independent researchers.

18. Phelan-McDermid Syndrome International Registry: Lessons from Citizen Scientists

Megan O'Boyle[1], Geraldine Bliss[1], Vanessa Miller[2], Catalina Betancur[3]

[1]Phelan-McDermid Syndrome Foundation; [2]Patient Crossroads; [3]INSERM

Background

Phelan-McDermid Syndrome Foundation is 13-year-old nonprofit serving approximately 900 members worldwide diagnosed with Phelan-McDermid Syndrome (deletion 22q13).

In February 2010, members of the PMSF wrote a Strategic Plan for Science. One of the first initiatives embarked upon was developing a web-based patient registry. The Phelan-McDermid Syndrome International Registry (PMSIR) portal allows parents to submit genetic reports as well as their child's medical and developmental history. As of March 2013, 530 of the estimated 900 known diagnosed patients are registered in the PMSIR.

The registry also includes a portal where researchers can see de-identified patient information. Researchers can use the data to understand how PMS features relate to the size and position of the genetic anomaly. Also, PMSF can use the data to contact families about participating in clinical studies and trials based on variables specified by clinical investigators.

With over 200 questions about PMS features, the Foundation has created the most comprehensive database of information about PMS patients worldwide. Through their participation families are meaningfully engaged in the quest for a cure.

Lessons Learned

- Too many questions can be overwhelming.
- 84% of registrants prefer shorter, multiple specific surveys to fewer, longer surveys.
- Finding medical records can be a barrier to completing surveys.
- Professional curation is a must because genetic reports vary tremendously.
- Researchers are willing to accept parent-reported data from the registry.
- Families want to contribute to the registry. Participating makes them feel empowered.
- Caregivers skip questions when they can't remember answers, but skipped questions skew the value of the data for researchers.
- The data viewed by parents after the surveys are completed allows them to view and communicate their child's symptoms and how they compare with others affected by the syndrome.
- Parents have been able to use the data from PMSIR to improve their child's medical care, such as testing, treatment, etc. by communicating the results to physicians.
- Researchers are able to review the data of the registry to determine if there is sufficient evidence to continue with a study with this population.
- Researchers will save time and funding by not having to recruit, consent and survey families since, the data is made available to them through PMSIR researcher portal.

19. Duplication of a Single Gene within the FGS3 Region in a Patient with FG Syndrome

Jennifer N. Sanmann[1]; Lois J. Starr[2]; Bruce A. Buehler[2]; Bhavana J. Dave[1]; Diane L. Pickering[1]; Denae Golden[1]; Jadd M. Stevens[1]; Warren G. Sanger[1]

[1]Human Genetics Laboratories, University of Nebraska Medical Center and the Munroe-Meyer Institute for Genetics and Rehabilitation, Omaha, Nebraska, USA; [2]Genetic Medicine, University of Nebraska Medical Center and the Munroe-Meyer Institute for Genetics and Rehabilitation, Omaha, Nebraska, USA

FG syndrome was first described in 1974 and was initially characterized as a rare X-linked disorder occurring only in males with multiple congenital anomalies and intellectual disability. Since the initial description, this diagnosis has been expanded to include females, which has broadened the potential mode of inheritance (e.g., autosomal, pseudoautosomal) in at least a subset of affected families. Patients with FG syndrome most commonly have mutations in the mediator complex subunit 12 (*MED12*) at Xq13. However, genetic studies have identified likely pathogenic aberrations in other regions of the X chromosome in some of the patients diagnosed with FG syndrome without identification of an associated genetic change in *MED12*. Specifically, mutations in filamin A (FLNA) at Xq28 (FGS2), and in calcium/calmodulin-dependent serine protein kinase (*CASK*) at Xp12.4-p12.3 (FGS4) have been observed in patients with FG syndrome. Linkage was shown to a third locus at Xp22.3 in two families with FG syndrome; however, no target gene has been identified to date. Currently, this region (Xp22.3→Xpter) is characterized as FGS3.

We report a male patient diagnosed with FG syndrome by two clinical geneticists independently. His findings included macrocephaly, a tall forehead, an anterior hair whorl, down-slanting palpebral fissures, deep-set eyes, small ears, slender hands/feet, and hypotonia. MRI of the brain showed agenesis of the corpus collosum. He also had intellectual disability, expressive speech delay, chronic constipation, temper tantrums, and head-banging. This patient's genetic work-up included normal *MED12* sequence analysis, normal male (46,XY) karyotype, normal chromosome 15 methylation studies, and normal Fragile X testing (less than 45 CGG nucleotide repeats).. However, array comparative genomic hybridization identified an approximately 147 kb duplication in the pseudoautosomal region 1 (PAR1) at Xp22.33/Yp12.3. This duplication is restricted to the short stature homeobox (*SHOX*) and includes all but exon 6b of this gene's coding region. This duplication is within the FGS3 region, which has been linked to two families with FG syndrome, for which no causative gene has previously been identified. Although certainly a heterogenic syndrome in nature, we propose that duplications of *SHOX* in the PAR1 could be the underlying genetic etiology for a subset of these patients and that microarray should be included in the testing regimen for FG syndrome.

20. A software solution for reviewing, visualizing, and annotating sequence variants alongside copy number changes

Soheil Shams[1], Zhiwei Che[1], Louis Culot[1], Shalini Verma[1], Andrea O'Hara[1] , Raja Keshavan[1]

[1]BioDiscovery, Inc. 5155 Rosecrans Ave, Hawthorne, CA 90250

Sequence variants (specifically point mutations, insertions, deletions, and inversions) can be penetrant for constitutional diseases as well as cancer, and should be considered alongside larger copy number variation. In Smith-Magenis syndrome, for example, the majority of cases are due to deletions of 17p11.2; but some cases have a mutation in RAI1 without a deletion. Since genome-wide copy number is typically arrived at through aCGH or SNP array analysis, and sequence variation through other methods, such as Next Generation Sequencing (NGS), an integration of the two data sources provides a more comprehensive picture of the genome than would otherwise be possible. We present here a software system for interpretation of these events simultaneously, presentation of integrated results for cytogenetics reporting (including the ability to create visualizations focused on suspected pathogenic events), case query of these events against history of previously analyzed samples, along with several actual cases.

21. The DECIPHER Project: Integrating sequence and structural variation

G.J.Swaminathan[1], E.Bragin[1], E.A.Chatzimichali[1], A.Paul Bevan[1], M.E.Hurles[1], H.V.Firth[1]

[1]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA. United Kingdom.
E-Mail: gs9@sanger.ac.uk

DECIPHER (https://decipher.sanger.ac.uk) [1] was established in 2004 as a web-accessible database to aid the interpretation of sub-microscopic chromosomal imbalance. Since its inception DECIPHER has gone on to become a powerful clinical and research tool for the deposition, analysis and interpretation of patient genomic microarray data and associated phenotypes. Contributing to DECIPHER is a consortium of over 250 academic departments of Genetic medicine comprising over 1500 clinical geneticists and cytogeneticists. Linked anonymous genomic and phenotypic data uploaded via a secure interface allow the confidential analysis and interpretation of patient phenotype-linked genotype data prior to sharing with other users following informed consent. This allows opportunities for contact and collaboration between consortium members or DECIPHER-facilitated contact between external users and consortium members. As part of its commitment to further medical and scientific knowledge about chromosomal microdeletions/duplications and facilitate research genes that affect human health, DECIPHER also makes available anonymised consented data to bona fide researchers. These collaborative efforts have been instrumental in the identification of many syndromes like 17q21.3, 14q11.2, 19q13.11 etc.

With over 22000 patient data records, DECIPHER occupies a strong position for genomic research and clinical diagnosis. In order to continue to provide an invaluable resource for bioinformatics and clinical research, we have recently extended DECIPHER to encompass all forms of genomic variation, including SNPs, InDels and CNVs. We have also developed new interactive graphical and interpretive bioinformatics tools to aid interpretation by facilitating simultaneous comparison of sequence and copy-number variation against both pathogenic resources (HGMD, LSDBs, ISCA) as well as population resources (dbSNP, dbVAR, Consensus CNV data). We present an update of all recent developments in DECIPHER and discuss our experience of the challenges and opportunities of bringing together sequence and structural variation.

Reference

1. Swaminathan, G.J. *et al.* (2012). DECIPHER: web-based, community resource for clinical interpretation of rare variants in developmental disorders. ***Hum. Mol. Genet.*** **21**, (R37-44).

22. Heterozygous deletion of *CHL1* gene: detailed array-CGH and clinical characterization of a new case and review of the literature. A new emerging syndrome?

Elisa Tassano[a], Roberta Biancheri[b], Laura Denegri[b], Simona Porta[a], Giorgio Gimelli[a], Cristina Cuoco[a]

[a]Laboratorio di Citogenetica, Istituto G.Gaslini, 5, 16147 Genova, Italy. [b]U.O.Neuropsichiatria infantile, Istituto G.Gaslini, 5, 16147 Genova, Italy.

Array-CGH has led to the identification of new syndromes by facilitating a ''reverse'' dysmorphology approach, in contrast to the previous ''phenotype first'' approach. In the present study, we report on a six-year-old boy with microcephaly, short stature, strabismus, mild mental retardation, learning and language delay and visuo-perceptual organization deficiency. Karyotype of the patient and his parents was normal. Array-CGH, performed with Human CGH Microarray Kit 180K, revealed a submicroscopic 3p26.3 terminal deletion inherited from his normal mother, including only the Close Homologue of L1 (*CHL1*) gene. Additionally, a 209.2 Kb duplication on chromosome 21 at band q22.3 has been identified in the patient and his father. Until now, only another three patients presenting heterozygous deletion of chromosome 3 at band p26.3, including only the *CHL1* gene have been described (Cuoco C. et al., 2011; Pohjola P. et al 2010). Comparing all these *CHL1* gene deletion cases, we suggest a common phenotype with verbal function developmental delay. Since the common deleted region comprised only the *CHL1* gene, this latter could be considered to be responsible for speech and cognitive developmental delay. *CHL1* gene is highly expressed in the central and peripheral nervous systems. It codes for neural cell adhesion molecules playing an important role in the building and functioning of the brain, because cell migration, axonal growth, fasciculation, synaptogenesis, and synaptic remodeling require cell-to-cell and cell-to-matrix interactions (Frints SG et al. 2003). Moreover, it is known that deficit in *CHL1* gene may cause damaging in cognitive processes. Studies carried out by Pratte et al., (2009) showed that *Chl1*-deficient mice reacted differently to spatial and object novelty. The homozygously deleted mice presented severe impairment of the capacity to react to both spatial and non-spatial novelty, while the heterozygously deleted mice were restricted in their ability to detect spatial changes maintaining their ability to discriminate novel objects. Similarly, our patient and some of the other reported demonstrated deficit in visual-perceptual organization. The concomitance of the genetic and phenotypic alterations could be a good evidence of a new emerging syndrome associated with the deletion of *CHL1* gene alone, although the identification of new cases is required.

23. Interstitial deletion 14q31.1q31.3 transmitted from the mother to her daughter both with features of hemifacial microsomia

Gimelli S[a], Cuoco C[b], Ronchetto P[b], Gimelli G[b], Tassano E[b*]

[a]Service of Genetic Medicine - University Hospitals of Geneva - Geneva - Switzerland
[b]Laboratory of Cytogenetics - G. Gaslini Institute - Genoa - Italy

Hemifacial microsomia (HFM) is defined as a condition affecting primarily aural, oral, and mandibular development. It belongs to the group of the oculo-auriculo-vertebral spectrum (OAVS). The principal features include facial asymmetry, secondary to maxillary and mandibular hypoplasia, underdevelopment of the external ear (microtia) often associated with preauricular skin tags or pits, and conductive hearing loss. Skeletal alterations mainly consist of vertebral anomalies (Gorlin RJ et al.2001). Most cases are sporadic, but there are rare familial cases that exhibit an autosomal dominant inheritance. Expression varies within families. Several chromosomal abnormalities have been associated with OAVS syndrome, but no recurrent chromosomal abnormalities were identified. Here we report a 13-months girl with an unilateral right-sided microtia, and mild facial asymmetry. Whole tympanic cavity was reduced in volume with dysplasia of malleo-incudal and stapes-incudal articulations. Her mother showed a preauricular skin tag on the right. A deletion of the long arm of a chromosome 14 was suspected by cytogenetic analysis. Array-CGH analysis, performed with Array CGH Kit 244B, identified a ~7.9 Mb deleted region on 14(q31.1-q31.3) inherited from her mother. The deleted region includes only seven genes, an open reading frame (*C14orf145*) and 5 Mb of "gene desert". In our opinion, among the seven the most striking is the gene *FLRT2*. This gene is a member of the *FLRT* gene family, which encodes for transmembrane proteins involved in cell adhesion and in the interaction with other proteins at the cell surface regulating gene expression in the cell. Expression studies of *FLRT2* demonstrated its role in the development of the craniofacial region during mouse embryogenesis (Haines BP et al. 2006; Gong et al.2009). Also the 5 Mb "desert gene" region between *SEL1L* and *FLRT2* genes might have some influence on the phenotype. In fact, stable "gene deserts" appear to harbor multiple distant regulatory elements physically linked to their neighboring genes (Ovcharenko I et al. 2005; Dermitzakis ET et al.2005; Baira et al.2008).

In conclusion, the phenotypic features of our proposita and her mother are strongly associated to the haploinsufficiency of some of the genes included in the deleted region. However, we cannot exclude a position effect on genes flanking the deletion breakpoints and/or a possible effect due to the lack of a copy of the "gene desert" region.

24. Parental balanced paracentric and pericentric inversions can result in inherited interstitial deletions and/or duplications: three families with unexpected recombinant chromosomes in phenotypically abnormal offspring

Jill C Taylor[1], Jane Watt[1], Monique Robertson[1], Vivienne Bickley[1], Kit Doudney[1], Alexa Kidd[1], Kate Gibson[2], Peter George[1]

[1]Cytogenetics and Molecular Pathology, Canterbury Health Laboratories, Christchurch, NZ; [2]Genetic Health Service NZ, South Island Hub, Christchurch, NZ

The offspring of an individual with a balanced pericentric or paracentric inversion may have an abnormal karyotype due to meiotic recombination and the formation of an unbalanced recombinant chromosome. The two typical recombinant chromosomes from a pericentric inversion are usually referred to as duplication/deletion recombinants. The terminal chromosome segments that are not involved in the inversion are duplicated and deleted. Rarely, a paracentric inversion can give rise to an acentric recombinant chromosome or a dicentric recombinant chromosome. A familial inversion is most often detected following a referral for infertility. Inversions may cause infertility by impairing meiosis (specifically spermatogenesis) or by early pregnancy loss due to a nonviable conceptus with significant genomic imbalance. Therefore an inversion carrier couple are generally given a low risk of live born phenotypically abnormal offspring.

However, the routine use of aCGH is beginning to unveil unusual inherited recombinant chromosomes that challenge the traditional theories. We describe four different examples of an unexpected recombinant chromosome derived from an apparently balanced parental inversion. In all three families, the proband was referred for microarray testing to investigate a developmental delay with congenital anomaly and/or dysmorphism. An interstitial deletion (two cases) and a complex interstitial deletion and duplication rearrangement (2 cases) were detected and subsequent parental studies revealed that the genomic imbalance was derived from a parental inversion. Standard G-band chromosome analysis and FISH studies were essential in the further characterisation of the recombinant and inversion chromosomes. We discuss potential mechanisms of formation, genotype-phenotype correlations and recurrence risks. This collection of cases indicates that the frequency of familial inversions resulting in interstitial genomic imbalance in offspring may be more common than traditional cytogenetic data has suggested.

25. SureCall: Identification and Classification of Mutations from Agilent's HaloPlex Sequencing Data on your Laptop

Ashutosh[1], Jayati Ghosh[1], Christian Le Cocq[1], Anniek De Witte[1], Stephanie Fulmer-Smentek[1], Olof Ericsson[1], Peter Webb[1]

[1]Genomics Solutions Division, Agilent Technologies, Santa Clara, California

As NGS transitions to use in clinical research for mutation detection and discovery, it is essential to have an efficient workflow from sample prep to data analysis. Agilent's HaloPlex technology is ideally suited to sample prep for deep sequencing of small panels of genes (up to 5 Mb), as is required for studies of inherited disorders or cancer. Desktop sequencers address the middle of the workflow. Now, Agilent's SureCall software, which runs on standard laptops or desktops, addresses the final critical steps by providing an easy-to-use analysis tool incorporating widely accepted open source libraries and algorithms.

Analysis in SureCall begins with data from Illumina HiSeq/MiSeq or Ion Torrent sequencing of genomic DNA enriched with a custom or catalog HaloPlex kit. After trimming adaptors, reads are aligned using BWA or TMAP. SAMTools recalibrates base call quality scores and performs local realignment. SAMTools also calls mutations and assesses the significance of the mutation calls. Each mutation is then evaluated based on its location, if it results in an amino acid change, its effect on protein function (SIFT), and its impact on protein structure and function (PolyPhen-2). Information regarding the mutation is then aggregated from public sources, including NCBI, COSMIC, and Pubmed. A proprietary mutation classifier then evaluates the significance of mutations following standard (e.g., ACMG) or customized guidelines. The user can then review each call using the built in genome viewer to show raw data and confidence measures, and links to external databases such as OMIM or dbVar. Any call made by the automatic workflow can be overridden or suppressed by the user.

We will present the analysis of eight datasets from samples prepared using three different HaloPlex research panels (Cardiac Disease; Noonan Spectrum Disorder; and Collagen Tissue Disease) and sequenced on an Illumina MiSeq. Analysis in SureCall identifies previously validated mutation calls and identifies several new calls missed in previous analyses.

Agilent's SureCall application, available at no cost to Agilent target enrichment customers, addresses all the needs of HaloPlex users for analysis, visualization, contextualization and summarization of their research results. This easy-to-use tool performs all analysis steps including alignment, mutation detection, mutation categorization, visualization, linkage to external databases, and guided ranking of mutations.

26. A 5 kb deletion of maternal *MEG3-DMR* causing a typical UPD14 clinical presentation

Dongbin Xu[1], Corbin Schwanke[1], Jonathan Schoof[1], Katherine Berry[1], Tessa Pitman[1], James Reynolds[1], Samuel Yang[1], John Johnson[1]

[1]Medical Genetics, Shodair Children's Hospital, Helena, MT.

A boy born at 34 weeks gestation presented with multiple congenital anomalies, including prenatal polyhydramnios, bell-shaped thorax, neonatal pneumothorax, hypotonia, abdominal distention, short neck, protruding philtrum, and other anomalies. Postnatal growth delay, macrocephaly, umbilical and inguinal herniae, scoliosis and dysphagia developed later. A normal male karyotype was detected. Given a clinical presentation consistent with paternal uniparental disomy 14 (patUPD14), including "handlebar" ribs on chest X ray, a family study was pursued, but showed normal inheritance for chromosome 14. Chromosomal microarray results were normal, with no CNVs greater than 40kb generally, or as low as 10 kb adjacent to genes, deemed to be significant. At that point, we decided to investigate imprinting of *Maternally Expressed Gene 3* (*MEG3*) at 14q32. The maternal *MEG3* allele is hypomethylated and remains active, while the paternal allele is oppositely regulated. Methylation analysis showed abnormal hypermethylation of *MEG3*. Re-analysis of the microarray indicated presence of a 5 kb deletion within the differentially methylated region (DMR) at the 5' end of *MEG3.* This deletion was confirmed by FISH analysis. Individuals with loss of the maternal copy of *MEG3* show typical paternal UPD14 findings. The microdeletion of the maternal *MEG3* allele in this patient caused a paternal UPD14 phenotype. The deletion was found in the patient's mother but not in his maternal grandfather or maternal aunt. This familial case represents a typical inheritance pattern of imprinting defects: the mother was unaffected by deletion of the normally silent paternal allele, but her son was affected due to loss of an active maternal allele. More importantly, this case highlights the necessity of taking all disease relevant genes into consideration when performing array data analysis. We also suggest that retrospective analysis of previous normal array results following clinical updates might assist in making a diagnosis of unresolved clinical cases.