Diagnostic Exome Sequencing Provides a Diagnosis for Half of Neonatal Patients: A retrospective analysis of detection rates and phenotypes among neonates.

Layla Shahmirzadi[1], Kelly Farwell[1], Tehila Ness[1], Stephanie Gandomi[1], Dima El-Khechen[1], Zöe Powis[1], Christina Alamillo[1], Kelly Radtke[1], Deepali Shinde[1], Christian Gund[1], Vruti Mehta[1], Marilyn Tsang[1], Katherine Helbig[1], Ira Lu[1], Cameron Mroske[1], Ruth Baxter[1], Brigette Tippin Davis[1], Elizabeth C. Chao[1,2], Sha Tang[1]

[1]Ambry Genetics, Aliso Viejo, CA, 92656 [2]Department of Pediatrics, University of California, Irvine, Irvine, CA 92697

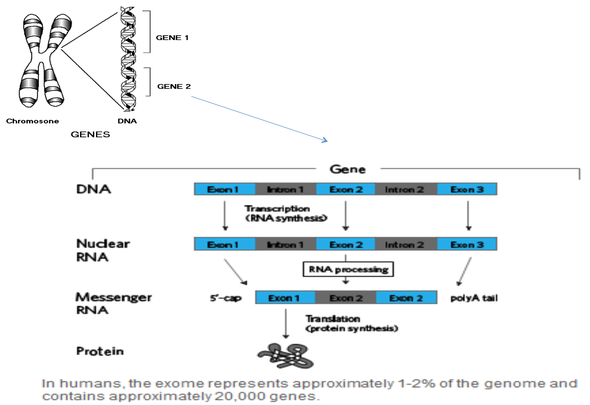
BACKGROUND

- Exome sequencing is the simultaneous sequencing and analysis of all of the exons (or protein coding) regions of the genome **(Figure 1)**.
- Analysis involves interpretation of the genes and the alterations. The genes involved include both characterized and novel genes.
  - Characterized genes: associated with human Mendelian disease
  - Novel genes: not associated with human disease.
- Diagnostic exome sequencing (DES) has been instrumental in establishing a molecular diagnosis in patients with previously undiagnosed disease, enhancing genetic counseling and aiding in clinical management.
- The aim of this study was to observe the application of DES in a neonatal population.
  - Neonatal: defined in this study as any patient presenting with symptoms prenatally or within the first 4 weeks of life and submitted for diagnostic exome sequencing at 4 months of age or less.

METHODS

- **Patients/study population:** Genomic deoxyribonucleic acid (gDNA) was isolated from whole blood from the first 500 probands and relatives referred to Ambry Genetics (Aliso Viejo, CA) for diagnostic exome sequencing (DES). Informed consent was obtained from all family members involved in the testing process.
- **Whole exome sequencing:** Samples were prepared using the SureSelect Target Enrichment System (Agilent Technologies, Santa Clara, CA) or SeqCap EZ VCRome 2.0 (Roche NimbleGen, Madison, WI). The enriched exome libraries were sequenced using paired-end, 100-cycle chemistry on the Illumina HiSeq 2000 (Illumina, San Diego, CA).
- **Characterized and Disease-causing (ChAD) and novel gene interanal databases**: The ChAD gene database was curated on a weekly basis to include genes currently known to be responsible for causing human disease. The ChAD database included genes which are associated with syndromes listed in the Human Gene Mutation Database (HGMD) (Stenson, 2009) and the Online Mendelian Inheritance in Man (OMIM) database. Novel genes were defined as those not known to underlie a Mendelian condition at the time of data analysis. Any RefSeq gene not included in the ChAD database was included in the novel gene database.
- **Bioinformatics annotation, filtering of variants, and Family history Inheritance-based Detection (FIND):** HGMD, OMIM, the Single Nucleotide Polymorphism database (dbSNP) (Sherry, 2001), 1000 Genomes, HapMap data (International HapMap, 2003) and online search engines (e.g., PubMed) were used to search for previously described gene mutations and polymorphisms. Stepwise filtering included the removal of common SNPs, intergenic and 3'/5' UTR variants, non-splice-related intronic variants, and lastly synonymous variants. Variants were then filtered further based on family history and possible inheritance models using the informatics program "FIND" (Family history Inheritance-based Detection).
- **Personalized Medical Review with Enhanced and Comprehensive Assessment (PRECISE) of potentially causal variants:** Each candidate mutation was assessed by a molecular geneticist to identify the most likely causative mutation(s) using the propriatary "PRECISE" (Personalized Medical Review with Enhanced and Comprehensive Assessment) analysis method. In brief, interpretive filtering was based on the deleterious nature of the candidate alterations, literature search, and analysis of the relevance of the candidate genes' function in relation to the patient's phenotype. Most candidate alterations undergo Sanger sequencing confirmation and familial co-segregation analysis.
- **Data curation:** Data regarding the primary exome test ordered, patient age, diagnosis and/or clinical description, and exome sequencing results were also curated. Data on results included gene(s), alteration(s), gene category (novel, characterized), alteration interpretation (pathogenic, likely pathogenic, uncertain, likely benign), and clinical overlap of gene-association and patient phenotype (positive, uncertain, partial).

FIGURE 1. DIAGNOSTIC EXOME SEQUENCING



In humans, the exome represents approximately 1-2% of the genome and contains approximately 20,000 genes.

FIGURE 3. COMMON INDICATIONS FOR TESTING

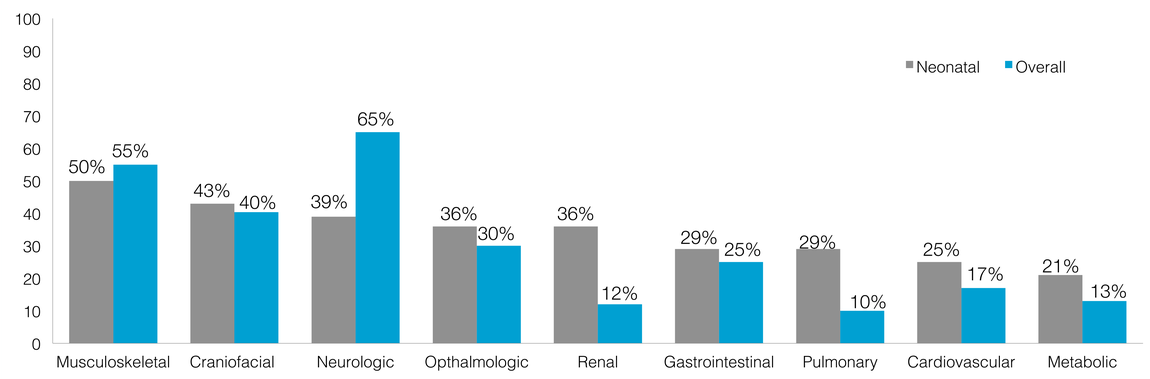


FIGURE 2. PATIENT POPULATION

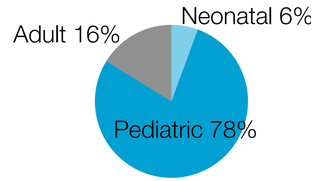


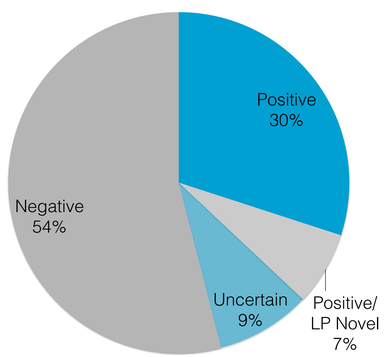
TABLE 1. SELECTED DIAGNOSES

GENE	SYNDROME
DNC2H1	Asphyxiating thoracic dystrophy 3 (ATD3) (aka Jeune Syndrome) (MIM_208500)
ACAT1	Acetoacetyl-CoA-thiolase deficiency (Beta-ketothiolase deficiency) (MIM_203750)
COL4A5	X-linked Alport Syndrome (MIM_301050)
LAS1L	NOVEL GENE
ACTG2	Visceral myopathy (MIM_155310) (recurrent)
AKR1D1	Bile acid synthesis defect, congenital (MIM_235555)
ALG1	Congenital disorder of glycosylation, type 1k (CDG 1k) (MIM_608540)
SF3B4	Acrofacial dysostoses, Nager syndrome (MIM_154400)
KCNQ2	Benign familial neonatal convulsions (BFNC) (MIM_121200)
GBE1	Glycogen storage disease (MIM_232500)
OFD1	Oral-facial-digital syndrome 1 (MIM_311200)
NPHP3	NPHP3 related nephronophthisis syndrome (MIM_608002)
ITSN1 and *DPYSL2*	NOVEL GENES

RESULTS

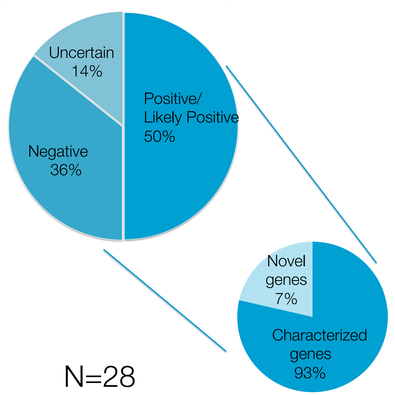
- A retrospective analysis of the first 500 patients submitted for DES at one laboratory revealed that the majority of cases were pediatric patients under the age of 18 (78%) **(Figure 2)**.
- 28 of the 500 cases (6%) included patients from the neonatal period, including 2 cases of fetal demise.
- The nine most common indications for testing among this neonatal cohort included musculoskeletal/structural (50%), craniofacial (43%), neurologic (39%), ophthalmologic (36%), renal (36%), gastrointestinal (29%), pulmonary (29%), cardiovascular (25%) and metabolic (21%) findings **(Figure 3)**.
- Musculoskeletal/structural features were the most common indication for testing among the neonatal cohort (50%), while neurologic findings were the most common indication among all the 500 patients analyzed (65%) **(Figure 3)**.
- Overall, detection rates were higher among the neonatal patients (50%) than among all patients undergoing DES (37%) **(Figures 4 & 5)**.
- Among the neonatal findings, 93% of positive/likely positive findings were in characterized genes while 7% were findings in novel genes **(Figure 5 & Table 1)**.

FIGURE 4. OVERALL DETECTION RATES



N=500

FIGURE 5. NICU DETECTION RATES



N=28

TAKE-HOME POINTS

- When compared to previously reported diagnostic rates among all patients undergoing whole exome sequencing, neonatal patients had up to 35% higher detection rates.
- These findings indicate an increased detection of Mendelian disease among neonatal patients, likely correlating with congenital or early-onset severe phenotypes.
- This data highlights the power of DES, particularly during the neonatal period, to help provide a diagnosis for patients with complex or nonspecific findings and to help guide clinical management and eliminate unnecessary treatments and testing during the neonatal period.

REFERENCES

- 1000 genomes database: A map of human genome variation from population-scale sequencing. *Nature* **467**,:1061-1073.
- Gonzalez KF, Shahmirzadi L, El-Khechen D, Powis Z, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from the 500 unselected families with undiagnosed genetic conditions. (Accepted Manuscript)
- International HapMap (2003) The International HapMap Project. *Nature* **426**:789-796.
- Ku CS, Cooper DN, Polychronakos C, Naidoo N, Wu M, *et al.* (2012) Exome Sequencing: Dual role as a discovery and diagnostic tool. *Ann Neurol* 71: 5-14.
- Sherry ST, *et* al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**,:308-311.
- Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), World Wide Web URL: http://omim.org/
- Stenson PD, et al. (2009) The Human Gene Mutation Database: 2008 update. *Genome medicine* **1** :13.
- Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, et al. (2013) Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med* 369: 1502-1511.

15 Argonaut, Aliso Viejo, CA 92656 Toll Free 866 262 7943 Fax 949 900 5501 ambrygen.com