Diagnostic Exome Sequencing Establishes Molecular Diagnoses Among Patients with Gastrointestinal Disease

Shahmirzadi L[1] , Farwell K[1], El-Khechen D[1], Powis Z[1], Alamillo C[1] , Radtke K[1] , Shinde D[1], Mroske C[1], Baxter R[1], Tippin B[1], Gandomi SK[1], Chao EC[1,2], Tang S[1]
[1]Ambry Genetics, Aliso Viejo, CA, 92656
[2]Department of Pediatrics, University of California, Irvine, Irvine, CA 92697

BACKGROUND

- A high diagnostic rate, the generation of clinically actionable data for patient care guidance, and dramatic cost-savings are just some of the features that make diagnostic exome sequencing (DES) ideally suited to be the standard of care for diagnostic medicine in the 21st century.
- This test is often indicated when a patient has a suspected genetic diagnosis and prior genetic testing has been uninformative, when there are multiple genes/diseases in the differential, or when the phenotype is not consistent with a known genetic disorder.
- DES has been instrumental in providing a diagnosis for patients with a broad range of indications.

METHODS

- **Patients/study population:** Genomic deoxyribonucleic acid (gDNA) was isolated from whole blood from 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 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 "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. COMMON INDICATIONS FOR TESTING

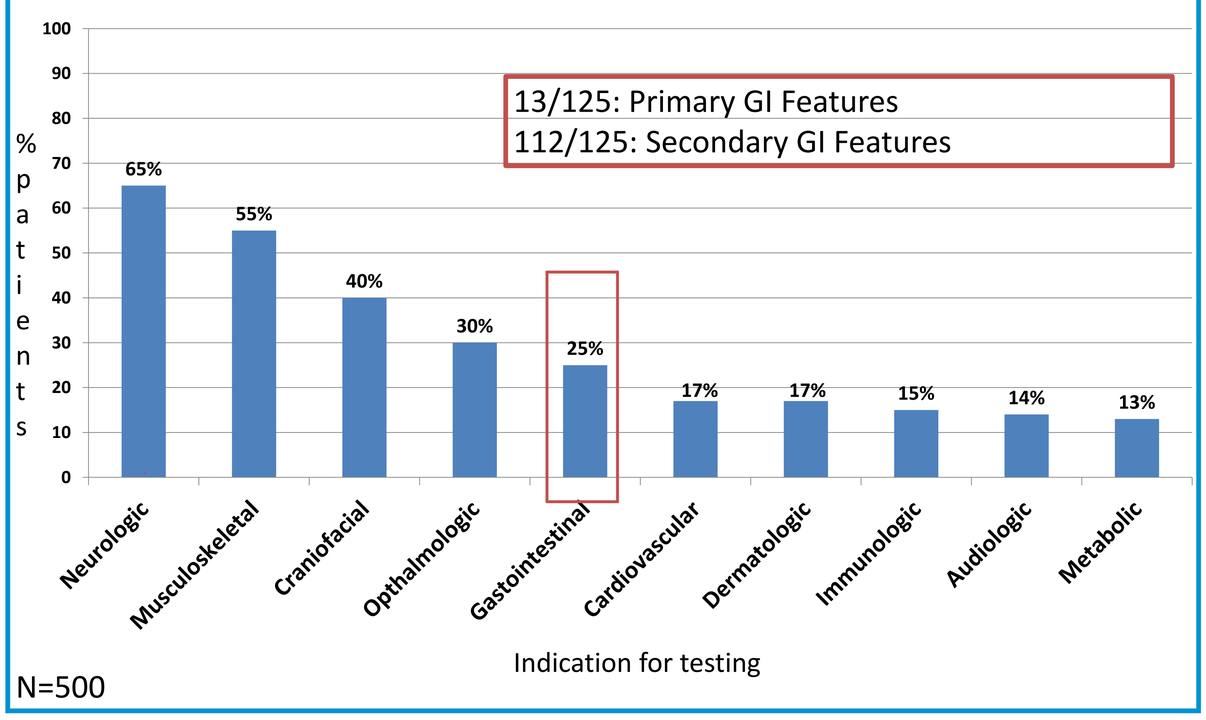


FIGURE 2. COMPARISON OF DIAGNOSTIC RATES AMONG ALL PATIENTS VS. PATIENTS WITH GI FINDING

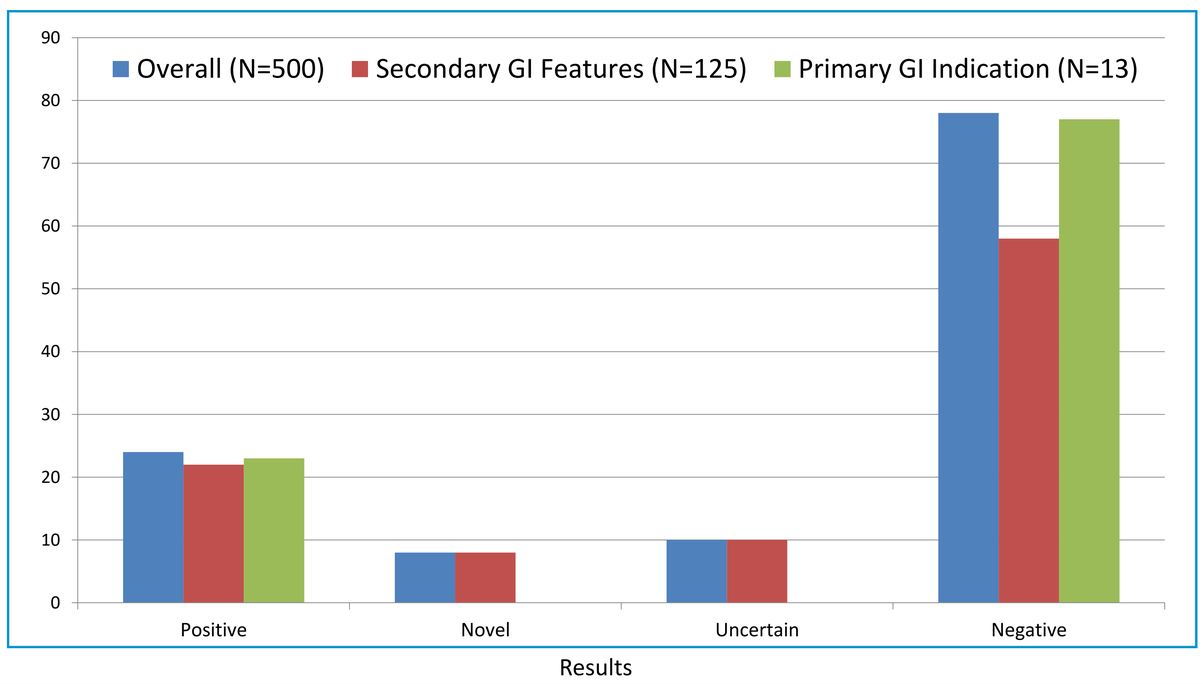
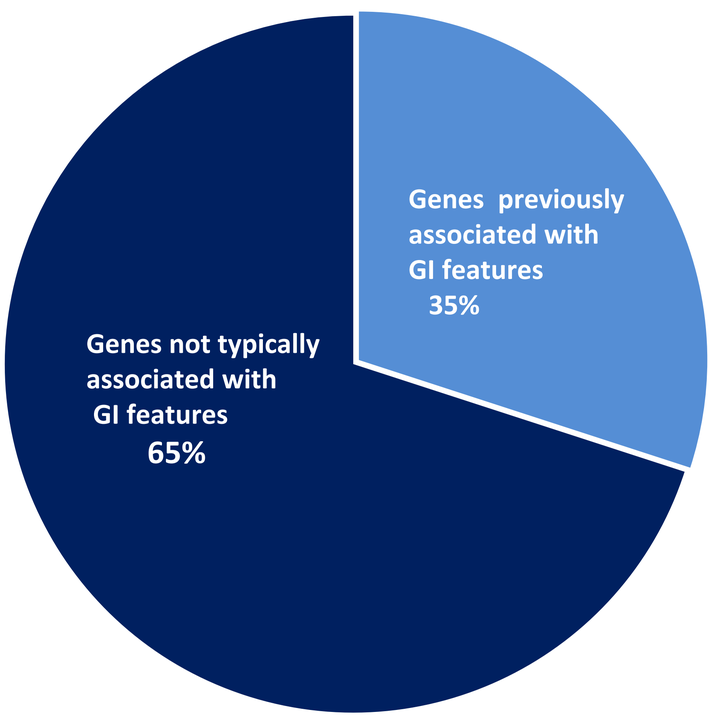


FIGURE 3. GENES IDENTIFIED WITH GI FEATURES AS PART OF CLINICAL SPECTRUM



Patients with GI disease and a positive finding (N=30)

RESULTS

- In a retrospective analysis of the first 500 reported patients who underwent DES at one laboratory , gastrointestinal (GI) features was one of the 10 most common indications for testing , where 25% (125/500) of patients presented with GI symptoms as one of their clinical features (**Figure 1**).
- 13/125 Patients presented with GI features as the primary indication for testing, while 112 patients presented with GI features among other primary concerns.
- A positive or likely positive finding was observed in 24% of all patients submitted for DES with both primary or secondary GI features. The diagnostic rates among all patients submitted were similar to those submitted for both primary GI features or secondary GI features **(Figure 2).**
- In thirteen cases, the primary presenting feature and indication for DES involved GI symptoms. Two patients presented with echogenic bowel, abdominal distention, bowel dysmotility and were each found to carry a heterozygous ***ACTG2*** mutation. A third patient primarily presented with omphalocele at birth and carried an ***OFD1*** mutation.
- 30 molecular diagnoses were established within well-characterized genes, while 5 cases identified pathogenic mutations in novel genes.
- Positive findings were identified in the following *characterized* genes: *TRPS1, ANKRD11* (2 cases), *TUBB4A* (2 cases), *ALG1, PTCHD1, SHANK3, STXBP1*, ***ACTG2*** (2 cases), *ARHGEF9, SCN1A, KAT6B, PKD1, FOXG1, SYNGAP1, NDUFA1, PCNT, SMC3, TMEM231,* ***OFD1****, ANO3, COL3A1, AMPD2, GARS, COL4A1, HMBS, NGLY1*.
- Positive findings in the *novel* genes *IL21R, MYH10, CACBA1E, ITSN1*, and *PURA* accounted for 14% .
- Of the characterized genes, only 35% (9) were described to have GI features as part of the clinical spectrum of disease **(Figure 3).**

NOTE: Genes in bold signify cases in which patients presented with GI features as the primary indication for testing.

TAKE-HOME POINTS

- Gastrointestinal features are among one the most common indications for diagnostic sequencing, however these features are often seen in combination with other primary indications for testing.
- Diagnostic rates among patients with GI features as both the primary or secondary indication for testing were similar to the diagnostic rates among all patients undergoing diagnostic exome sequencing.
- This data suggests an expansion of the clinical spectrum of disease among some genes typically not associated with gastrointestinal features. However, due to the high frequency of GI features in the pediatric population, some of these associations may be coincidental.

REFERENCES

- 1000 genomes database: A map of human genome variation from population-scale sequencing. *Nature* **467**,:1061-1073.
- International HapMap (2003) The International HapMap Project. *Nature* **426**:789-796.
- 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), [date]. World Wide Web URL: http://omim.org/
- Stenson PD, *et al*. (2009) The Human Gene Mutation Database: 2008 update. Genome medicine **1** :13.