Yield of Clinically Relevant Candidates in Family Genomes in the UK 100,000 Genomes Project Using the Fabric Genomics Platform

FABRIC GENOMICS™

Melanie Babcock[1], Charlene Son Rigby[1], Marco Falcioni[1], Andrew Guo[1], Josh Grigonis[1], Kyle Hart[1], Mark Yandell[2], and Martin G. Reese[1]. 1. mbabcock@fabricgenomics.com, Fabric Genomics™ Inc., Oakland, CA; 2. Department of Human Genetics, University of Utah, Salt Lake City, UT

ABSTRACT

The 100,000 Genomes Project, spearheaded by Genomics England (GeL), is a United Kingdom National Health Service sponsored study aimed at identifying disease-causing genetic variants in patients and families with rare genetic diseases and cancer using a whole genome sequencing (WGS) approach. For this study, clinical history was used to recruit patients into specific disease categories, each of which were associated with gene panels curated in the GeL PanelApp tool.

Fabric Genomics™, a clinical interpretation partner for the 100,000 Genomes Project, has analyzed over 1973 clinical cases using Opal™ Clinical, powered by Fabric Enterprise, a comprehensive platform for NGS data analysis and clinical reporting. The variant filtering and prioritization protocols utilized for case analysis include GeL's variant tiering methodology, ClinVar, and Fabric Genomics' proprietary variant and gene ranking algorithms VAAST (Variant Annotation, Analysis and Selection Tool) and Phevor (Phenotype Driven Variant Ontological Re-ranking Tool). We report results showing that by applying VAAST and Phevor we increase the clinical candidate yield compared to using the GeL tiering system alone. We identified candidate causal genes/variants in 49.8% of the cases. In 23.0% of these cases (9.4% overall) candidates were only obtained by using the VAAST/Phevor top 20 ranked genes/variants.

In a subset of 1191 cases, we reviewed the effects of providing parental genomes in the analysis and return rate of results. Interestingly, there was only small difference in clinical candidate yield between solo cases and true trios (proband and unaffected parents), but by further investigating how the causal candidates were determined, it showed an increase of candidates that were identified exclusively from VAAST/Phevor top 20 ranked genes in true trios compared to solos.

Interpretation Methodology

We have implemented a whole genome interpretation methodology within Opal™ Clinical that is comprehensive and rapid.

The workflow developed for GeL includes consideration of the following variants:

- Variants scored under different inheritance modes in the top 20 genes ranked by VAAST and Phevor, two sequentially applied proprietary algorithms: VAAST integrates sequence conservation, genetic consequence, and allele frequency in a probabilistic framework to identify disease-causing alleles, Phevor then combines HPO-based patient phenotype descriptions with the VAAST results to re-rank the variants

- Variants in phenotype matched panels curated in the GeL PanelApp tool and categorized using GeL's tiering methodology (https://bioinfo.extge.co.uk/crowdsourcing/PanelApp/): known pathogenic, protein truncating, and de novo protein alternating variants (Tier 1), and other protein altering variants (Tier 2)

- Variants with at least one ClinVar Pathogenic or Likely Pathogenic submission

The interpretation methodology concludes with an expert review by Fabric Genomics' Clinical Services team. Each case was interpreted by two primary reviewers. Identified candidates were scored using the ACMG variant interpretation guidelines.

The candidate genes/variants are further evaluated by the NHS Genomic Medicine Centre laboratories and these labs will determine which of those candidates are reported back to the patients.

GeL Interpretation Request → API Upload & Annotation → VAAST & Phevor Analysis → Fabric Genomics Clinical Services → Genomics England → GMC → Clinical Report

GeL workflow for receiving interpretation requests, processing samples and returning clinical reports to Genomic Medicine Centres
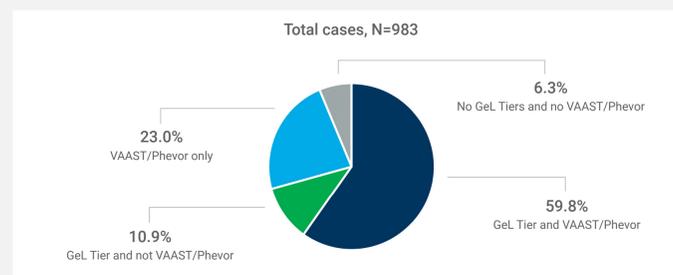
Fabric Genomics' Gene Ranking Algorithms: VAAST[1] and Phevor[2]

VAAST is an algorithm that was developed in collaboration with the University of Utah. Using VAAST for gene prioritization speeds diagnosis and improves diagnostic yield by providing a ranking of genes based on their likelihood to cause disease. Every variant is assessed for comparative functional impact on the protein product, conservation of the position across species, and the allele frequency.

Phevor re-ranks genes that have already been prioritized by VAAST by using the Human Phenotype Ontology (HPO) terms provided for the proband. Phevor starts by mapping phenotype terms to the Human Phenotype Ontology, Gene Ontology and other ontologies, then uses a unique network propagation approach to identify additional gene candidates. This process creates a ranked list of genes ordered by the specific phenotype provided. Phevor then combines this prioritized list of genes with the VAAST analysis to produce a combined ranking of candidate genes based on deleteriousness and the specific phenotype or phenotypes in question. This re-ranking allows rapid discovery of clinically relevant variants in genes related to the proband's phenotype, including genes not directly annotated to that phenotype.

Return of Results in 49.8% of 1973 Cases

Of the 1973 cases, we returned 49.8% with causal candidate genes/variants. Of cases with causal candidates, 59.8% had at least one tiered and one VAAST/Phevor Top 20, and 23.0% were identified exclusively from VAAST/Phevor Top 20 (9.4% improvement).

Total cases, N=983

- 6.3% No GeL Tiers and no VAAST/Phevor
- 23.0% VAAST/Phevor only
- 10.9% GeL Tier and not VAAST/Phevor
- 59.8% GeL Tier and VAAST/Phevor

Return of Results by Recruited Disease Category

Recruited Disease Categories are defined by the recruiting Genomic Medicine Centres (GMC). The rate of return of results was highest for Cardiovascular Disorders, Opthalmological Disorders, and Hearing and Ear Disorders.

Returned Results by Disease Category

- Returned Results
- No Solution
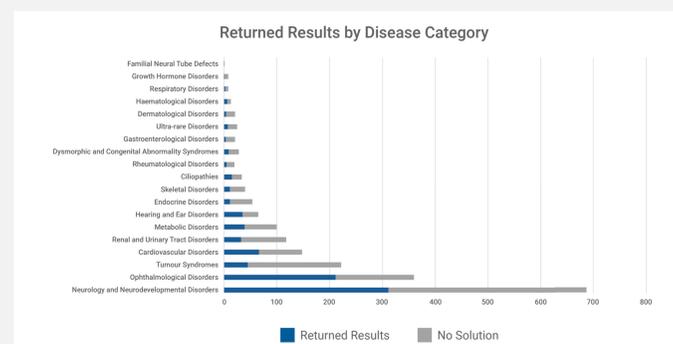
Return of Results Based on Family Structure

Of a subset of cases, 1191, we returned 38.6% with causal candidate genes/variants in proband only cases and 37.1% in parent-offspring trio cases (proband and 2 unaffected parents). When these results are broken down to the types of candidates, there is an increase of causal candidates that were identified exclusively from VAAST/Phevor Top 20 in true trio cases versus proband only.

- Parent-offspring trio
- Proband Only
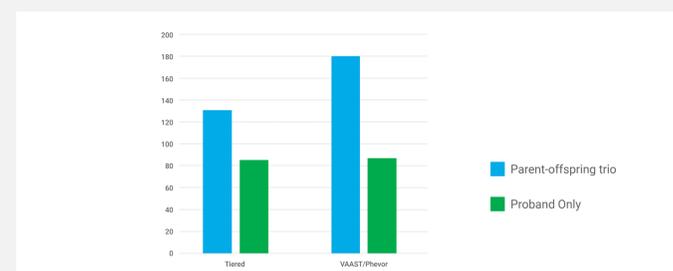
Example Cases Where Causal Candidates Were Only Found Using VAAST/Phevor

Main Recruited Disease Category	Condition	HPO/Phenotype	Panel Applied	Gene	Tier Status	VAAST Rank	Phevor Rank	ACMG Class	Family Style
Ophthalmological disorders	Noonan or Leopard	Retinal coloboma, visual impairment, retinal dystrophy, microphthalmia	Ocular coloboma	PTPN11	Not in the panel applied (present in another panel)	4	2	Pathogenic	True trio
Neurology and Neurodevelopmental Disorders	Intellectual developmental disorder with dysmorphic facies and ptosis	Global Developmental Delay, obesity, unilateral ptosis, short palpebral fissure, polyhydramnios	Intellectual Disability; RASopathies; Significant early-onset obesity +/- other endocrine features and short stature	BRPF1	Not in the panel applied (present in another panel)	2	3	Pathogenic	True trio
Ultra-rare disorders	Acampomelic campomelic dysplasia	Cleft palate, micrognathia, abnormality of the nervous system, small for gestational age, polyhydramnios, laryngomalacia, bilateral talipes equinovarus, gastroesophageal reflux, scoliosis, tracheomalacia, bronchomalacia, kyphosis, abnormality of the musculature, congenital septal defect, fixed elbow flexion, primum atrial septal defect, moderate intrauterine growth retardation, abnormality of limb bone	Intellectual Disability; Familial non syndromic congenital heart disease	SOX9	Not in the panel applied (present in another panel)	3	1	Pathogenic	True trio
Neurology and Neurodevelopmental Disorders	Mental retardation and microcephaly with pontine and cerebellar hypoplasia	Microcephaly, abnormality of the eye, autistic behavior, delayed speech and language development, intellectual disability, generalized hypotonia, small for gestational age, delayed gross motor development, proportionate short stature, unilateral strabismus	Intellectual Disability; Mitochondrial disorders, Undiagnosed metabolic disorders	CASK	No tier status	1	1	Pathogenic	True trio

CONCLUSIONS

Fabric Genomics' platform has provided GeL with potential causative candidates in 49.8% of cases. In the 1973 cases we present here, GeL's tiering system achieves a results return rate of 28.8%. By using our VAAST / Phevor algorithms we were able to increase the yield of candidate genes/variants by 23.0%, thus highlighting the complementary utility of the VAAST and Phevor algorithms and GeL's tier filtering methodology. There was a noticeable increased yield of causal candidates in true trios compared to solos that were identified only by VAAST/Phevor top 20. This further supports using our algorithms to help identify clinically relevant candidates for whole genome/exome analysis.

Fabric Genomics™ supports labs to maximize their diagnostic yield. The VAAST and Phevor ranking algorithms accelerate identification of disease-causing candidates. Fabric Genomics' clinical solutions are optimized to provide efficiency for variant scientist's time: the cases reported here each took less than three hours; however there are cases that took significantly less time. This time included both identification of candidate variants, and scoring using the ACMG 2015 Variant Interpretation Guidelines.

This makes Fabric Enterprise ideal for hard to solve rare genetic disease cases, and for large scale country projects.

1. Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. Rope et al. Am J Hum Genet. 2011 Jul 15;89(1):28-43.

2. Phevor combines multiple biomedical ontologies for accurate identification of disease-causing alleles in single individuals and small nuclear families. Singleton et al. Am J Hum Genet. 2014 Apr 3;94(4):599-610

FABRIC GENOMICS™ FORMERLY Omicia

Learn more at
www.fabricgenomics.com   info@fabricgenomics.com   510.595.0800