591

Integration of PHEVOR algorithm module to aid candidate gene identification in a clinical WES analysis pipeline: Prospective review across 74 consecutive cases

Wenjie Chen, PhD; Merry Ferre, MS, CGC; Merica Gellerman, MS, CGC; Ekaterina Bogdanova, PhD; Sabrina A. Gardner, PhD; Wei S. Kelly, PhD; Ildiko Thibodeau, PhD; Katelyn S. Weymouth, PhD; Ian King, PhD; Alecia Willis, PhD, FACMG; Narasimhan Nagan, PhD, FACMG
Integrated Genetics, Laboratory Corporation of America® Holdings, Westborough, MA and Raleigh, NC

Integrated GENETICS
LabCorp Specialty Testing Group

I. Abstract

Whole exome sequencing (WES) has been increasingly used to identify disease-causing alleles in novel and well-established genes to facilitate the clinical diagnosis of rare diseases. Tens of thousands of variants can be identified in any patient sample using WES; only one or a few of which are expected to be causative. A candidate gene prioritization strategy that ranks the underlying variants and the genes they affect, with those most likely related to the patient phenotype remains the corner stone for an unequivocal WES diagnosis. Omicia Opal Clinical™, a recently developed application, uses VAAST and PHEVOR algorithms to analyze WES data and prioritize variants and genes involved in disease. In this study, we report our experience with Omicia Opal Clinical™ to identify disease-causing variants across a set of 74 consecutive clinical WES cases. We identified at least one variant in an associated/novel possibly associated gene or a candidate gene in 43 out of the first 74 (58%) clinical WES cases analyzed. An additional set of 10 positive cases were obtained from our validation cohort for a total of 53 cases to assess the performance of the VAAST and PHEVOR algorithms. These 53 cases comprised a total of sixty-four identified genes spanning *de novo* (n=12), autosomal recessive (n=24), autosomal dominant (n=20), autosomal unknown (n=3), and X-linked (n=5) modes of inheritance. Ninety-eight percent (63/64) of the reported genes harboring the disease-associated variants were ranked by the VAAST-PHEVOR analysis. In the absence of HPO terms corresponding to the patient's phenotype, the percentage of genes that ranked within the top 20 was 66%. After incorporating HPO terms corresponding to the patient's phenotype in the VAAST-PHEVOR analysis, this figure increased to 83%. Furthermore, restricting the review to the Phenotype/Gene Association score in PHEVOR analysis improved the percentage of genes ranking within the top 20 to 88%, which is comparable to the ranking result obtained by the VarElect (free trial version, LifeMap Sciences) (89%) across the set of 53 identical cases. The impact on candidate gene ranking by PHEVOR analysis was more significant for singleton/duo cases than in trio cases. As a limitation, the VAAST-PHEVOR analysis does not rank genes with a missing 2nd variant in cases with possible recessive inheritance, requiring alternative approaches to rank these genes. In conclusion, the VAAST-PHEVOR analysis is an efficient adjunct for identification of disease-causing genes and variants in clinical whole exome sequencing analyses

II. Methods

Clinical Samples

Seventy-four clinical cases and ten validation cases were included in this study.

Whole Exome Sequencing and Sanger Confirmation

Genomic DNA is isolated from the provided specimens. The DNA samples are fragmented through sonication and exonic regions are enriched using Agilent SureSelect XT with custom content.
The enriched targeted DNA is sequenced on an Illumina HiSeq 2500 system. Sequence data is mapped and aligned to Human Genome Build GRCh37/hg19 using CLC Bio software. Variants included in the report are confirmed through targeted Sanger sequence analysis. This test was developed and its performance characteristics determined by LabCorp.

Data Interpretation and Reporting

Candidate genes/variants were selected based on variant type, inheritance modeling, presence in the literature, frequency in exome aggregation consortium (ExAC) population database, and phenotype correlation. Reported candidate genes were categorized as: disease genes related to phenotype (category 1), disease genes possibly related to phenotype (category 2), findings in strong candidate genes (category 3). VAAST and Phevor analysis were performed using Omicia Opal Clinical™. VarElect analysis was performed using the free trial version (http://varelect.genecards.org/). To compare the candidate gene ranking either PHEVOR or VarElect were applied to the output of VAAST analysis derived from Omicia Opal Clinical™ using identical HPO terms.

III. Results

Figure 1: A. The set of 74 consecutive clinical WES cases includes 55 trios (74%), 8 duos (11%), 11 singletons (15%). **B.** At least one pathogenic/likely pathogenic variant was identified in 15 cases (20%). At least one VUS variant was identified in 28 cases (38%)

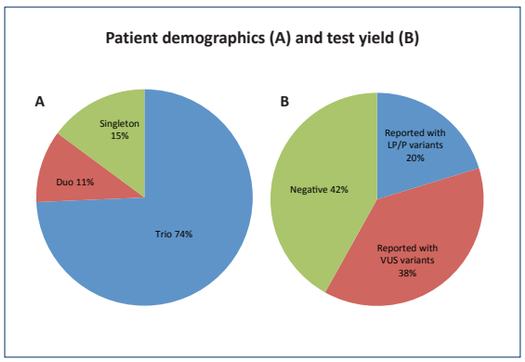

Table 1: Breakdown of test yield by number of reported genes per case

Total positive cases	43
Total reported genes across all positive cases	54
Cases with 1 reported gene	33
Cases with 2 reported genes	9
Cases with 3 reported genes	1

Figure 2: Ranking of 64 reported genes (43 clinical cases and 10 validation cases) by Opal VAAST, VAAST-Phevor, Phevor Phenotype/Gene Association (P/G), and VarElect.

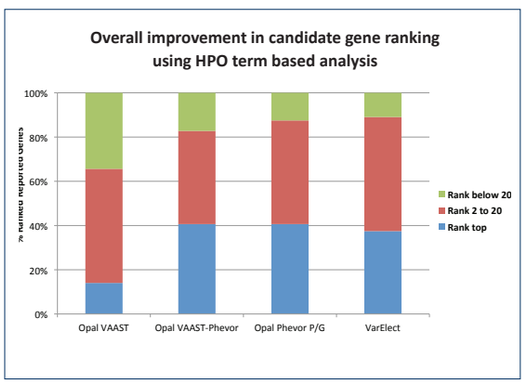

Figure 3: Ranking of 29 reported genes in trio cases and 35 reported genes in singleton/duo cases by Opal VAAST, VAAST-Phevor, Phevor Phenotype/Gene Association (P/G), and VarElect.
* p<0.05 when compared to candidate gene ranking using Opal VAAST.

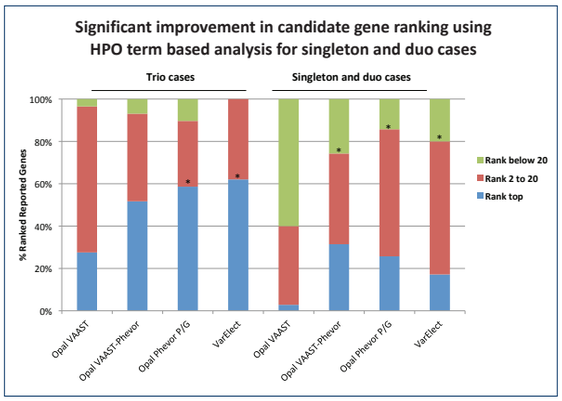

Figure 4: A. Category 1: disease genes related to phenotype; Category 2: disease genes possibly related to phenotype; Category 3: findings in strong candidate genes. **B.** P: pathogenic, LP: likely pathogenic, VUS: variant of unknown significance

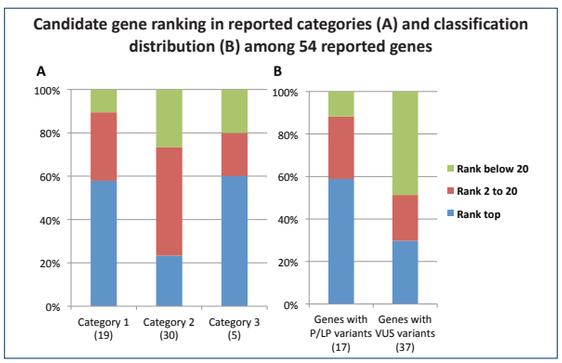

Table 2. Representative candidate genes

Gene	Total Genes in VAAST	Opal VAAST	Opal VAAST-Phevor	Opal P/G	VarElect
DYNC1H1	257	221	72	4	5
COL6A1	511	428	59	1	2
*DHCR7**	351	208	31	2	5
TTN	430	405	134	3	1
PTCH2	302	2	1	1	83
C5ORF42	33	5	31	31	1
C10ORF2	408	149	392	383	1
ADCY5	Not included in VAAST-Phevor due to low allele fraction (mosaicism)				

*The second allele is missing in DHCR7 gene. This gene was ranked in autosomal dominant mode in Opal VAAST and VAAST-Phevor.

IV. Conclusion

- HPO term based analysis significantly increases the percentage of top ranked candidate genes in all cases and the percentage of candidate genes ranked within the top 20 in singleton and duo cases.
- The VAAST-PHEVOR analysis does not rank candidate genes with a missing 2nd variant in cases with possible recessive inheritance. The alternative is to rank these genes in autosomal dominant mode.
- Overall, the VAAST-PHEVOR analysis is an efficient adjunct for identification of disease-causing genes and variants in clinical whole exome sequencing analyses.

LabCorp