Comprehensive Clinical Grade Whole Genome Sequencing Significantly Improves Diagnostic Yield in Sick Neonates And Pediatric Suspected of a Genetic Disorder

Madhuri Hegde[1], Ruby Liu [1,], C. Alexander Valencia[1], Leah Stansberry[1], Zeqiang Ma[1], and Alka Chaubey[1,]
[1]PerkinElmer Genomics, Duluth, GA

ABSTRACT

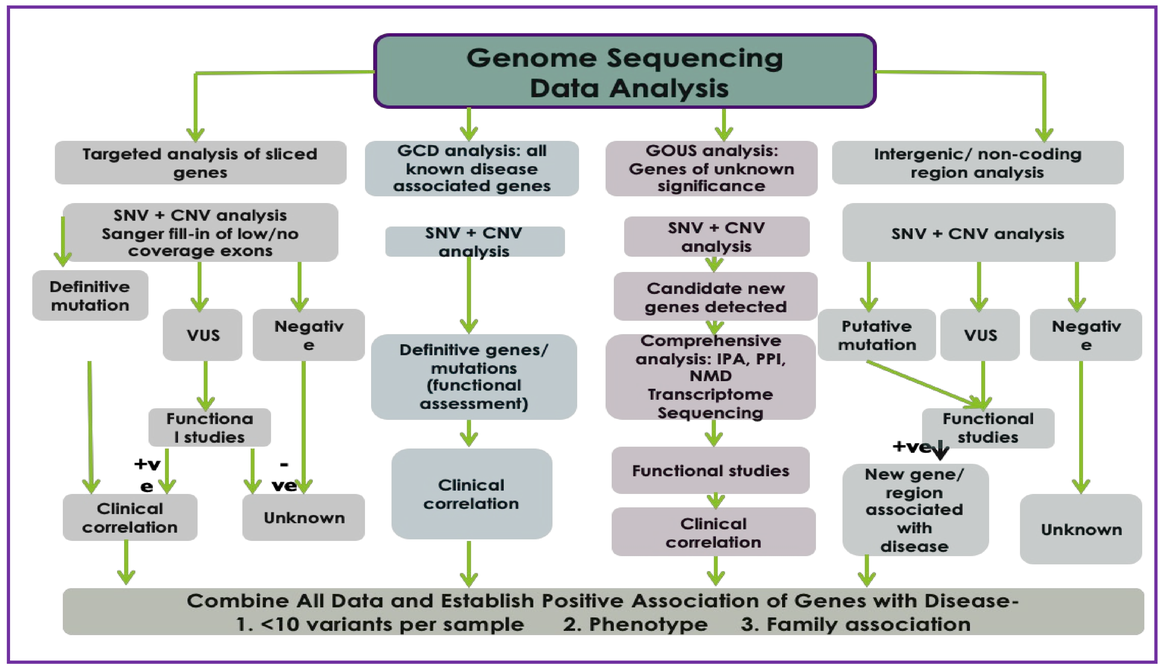
Whole-exome sequencing (WES) is now routinely used in clinical genetic testing and is optimized for the detection of rare and common genetic variants in humans. However, Whole genome sequencing (WGS) is becoming increasingly attractive as an alternative, due to its uniform coverage, ability to detect all types of variants and decreasing cost. We present here data from the first 150 cases referred to our laboratory for clinical grade WGS. 75% of the submitted cases had some form of previous cytogenetic and/ or molecular testing with either a negative result or a variant of unknown significance. Comprehensive clinical grade WGS has been validated in our laboratory at a depth of 30X and mitochondrial genome at a depth of 1000X, using a PCR free library preparation protocol followed by sequencing on the Illumina NovaSeq. Primary data processing is performed the Edico Dragen system, and bioinformatic analysis using our in house proprietary program ODIN (Ordered Data Interpretation Network). The WGS assay detects single nucleotide and copy number variation- CNV (genomic and intragenic) across the genome, along with mitochondrial genome. Data was parsed into categories: Genes causing disease (GCD: 5300 genes), Genes of Unknown Significance (GOUS), ACMG 59, known common and founder pathogenic changes, intragenic and intergenic variants with tagged variants which have been established to be diseases causing. In the 150 cases presented here, 96 cases were sent as singletons and 44 cases were sent at trios. 35 cases were from newborn in NICU or hospitalized and 115 were pediatric cases presenting to a clinical genetics clinic. Diagnostic findings were obtained in 47% of the cases, Variants of unknown significance in 33% and 20% were negative. The WGS assay not only provides a cost and time effective single molecular assay but most importantly the results in our cohort demonstrate the power of the WGS assay- 1. Detection of pathogenic variants previously missed on a gene panel: A c.669dupC duplication pathogenic variant as detected in the *PRRT2* gene which was missed on the previous gene panel which included the *PRRT2* gene in a child with epilepsy, 2. An intragenic duplication CNV spanning exons 3-14 in the *SETD5* gene was detected in child with Mild lateral and ventricular dilation suggestive of diffuse atrophy. This duplication was filtered from the previous microarray results due to low resolution. Confirmatory testing on the *SETD5* duplication was performed using qPCR and was shown to be de novo. 3. A promoter variant c.-106T>C was detected in the child with OTC which was not detected on previous testing. 4. In a 7 year old male, Developmental delay, Speech delay, Spastic paraplegia, Abnormal EEG sent to the lab as singleton , a *AP4M1* gene nonsense pathogenic homozygous variant p.R441X was detected. This variant was presented in a AOH region further demonstrating that this variant was homozygous thus ruling out the possibility of deletion on the opposite allele. Additional cases with interesting results included single exon deletions, mitochondrial pathogenic variant and tagged deep intronic variants. A second pathogenic variant was detected in five cases in which only a single pathogenic variant was reported in a gene related to phenotype. The average TAT in STAT cases was 5-10 days and in other cases 4-6 weeks. As we learn more about the non-coding regions of the genome we anticipate better clinical interpretation and thus more powerful re-analysis using WGS data in the near future. With cost of WGS reducing rapidly, it is time to strongly consider WGS as a first-tier molecular assay.

INTRODUCTIONS AND METHODS

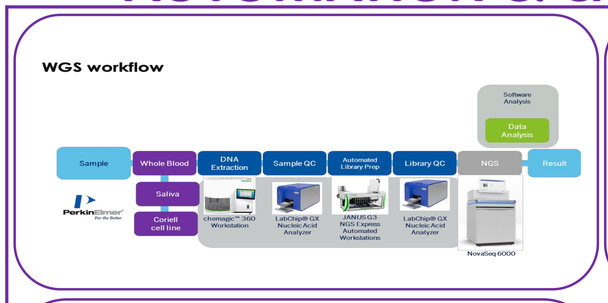
- **APPROACH: 30**X WGS was performed by using the KAPA HyperPlus PCR-free library construction kit and sequenced on Illumina NovaSeq™ 6000 (2 x 150 bp mode).
- **CLINICAL SAMPLES: USA, South America, Malaysia and India.**
- **AUTOMATION:** Unified automated workflow for high throughput and cost effective method for the 30X WGS assay
- **ANALYSIS and INTERPRETATION: PKIG proprietary ODIN software for SNV and** NxClinical 5.0 software (BioDiscovery, El Segundo, CA) was utilized for analysis, interpretation and reporting of CNVs and AOH. Multi Scale Reference algorithm is utilized by NxClinical 5.0.
- **30X WGS Data Quality in ODIN and NxClinical 5.0 software:**
  - ~90% of the human genome had coverage distribution between 30-40X depth
  - >1000-1250 million BAM reads were utilized for generation of log2 ratio and virtual copy number probes.
  - Mitochondrial genome depth 1000x-1500x
  - Low Coverage nucleotides/ exons ranged from 1-2%

RESULTS

- All clinical/diagnostic cases were subjected to 30x WGS assay (aligned to human reference genome hg19).
- 79 (53%) cases were classified with known pathogenic SNV / CNVs of clinical significance.
- 35 cases were NICU and 115 cases were pediatric.
- Pathogenic variants, SNV and CNV detected in 81 genes and two or more cases in *WDR45* and *KCNQ1* genes.
- **Dual Mendelian Diagnoses:** 2 cases with dual clinical diagnoses resulting from 2 pathogenic CNVs were identified.
- Mitochondrial pathogenic variant identified in one case.
- CNV identified in 4 cases- In addition, two cases of Trisomy 21.
- Two deep intronic pathogenic variants identified.



AUTOMATION & CASE EXAMPLES



Test Performed: Whole Genome Sequencing, Proband only

TEST RESULT SUMMARY

A hemizygous c.-106C>A *OTC* variant of unknown significance was identified in this sample. Clinical and biochemical correlation is required.

Diagnostic findings:

Gene OMIM	Associated Disease (Inheritance)	Exon/ Intron	DNA Change*	Protein Change	Zygosity	Classification
OTC 300461	Ornithine transcarbamylase deficiency (XLR)	5'-untranslated region	c.-106C>A	-	Hemizygous	Variant of Unknown Significance

XLR= X-linked recessive
* Sanger confirmation is pending.

TEST RESULT SUMMARY

Preliminary Report: A duplication of *SETD5* including exons 3-14 was identified in this sample. Clinical and biochemical correlation is required.

Diagnostic findings:

Gene OMIM	Associated Disease (Inheritance)	Location of Most 5' Abnormal Probe[a]	Location of Most 3' Abnormal Probe[a]	Classification
SETD5 615761	Mental retardation, autosomal dominant 23 (AD)	g.9470573 c.-50 (5'UTR)	g.9489001 c.1782+10 (intron 14)	Variant of Unknown Significance

AD=autosomal dominant
a. Probe location does not indicate precise breakpoint.

TEST RESULT SUMMARY

Pathogenic variant in the *PRRT2* gene was detected in this individual. Clinical and biochemical correlation is required.

DIAGNOSTIC FINDINGS RELATED TO PHENOTYPE

Phenotypic terms applied: Generalized seizure and infantile spasms

Gene OMIM	Associated Disease (Inheritance)	Exon/ Intron	DNA Change	Protein Change	Zygosity	Classification
PRRT2 614386	*PRRT2*-associated paroxysmal movement disorders (AD)	2	c.649dupC	-	Heterozygous	Pathogenic
KCNJ10 602208	SESAME syndrome (AR); Enlarged vestibular aqueduct (AR)	2	c.530A>G	p.E177G	Heterozygous	Variant of unknown significance

DISCUSSION/CONCLUSIONS

- **Diagnostic performance and clinical significance:** 30X assay has the potential to become a single diagnostic assay
- **Highest possible accurate breakpoint determination:** Approximate base pair resolution of deletions and duplications for microdeletions and microduplications and intragenic deletions/ duplications.
- **Both nuclear and mitochondrial genomes are covered**
- **Fast TAT and reduced cost:** Simple automated workflow reduces the TAT and multiplexing on the NovaSeq is extremely cost effective.
- **AOH detection**: AOH has been validated for various degrees of familial relationships
- **WGS as standard of care:** 30X WGS is an effective method for the diagnosis of mendelian diseases
- **Assay can be performed on dried blood spot card, saliva and whole blood**
- **Significant improvement in uniformity of data quality over exome sequencing and reduction in low coverage exons and gene.**
- **Significantly improved diagnostic yield and time to diagnosis**

Disclosure: All the authors are employees of Perkin Elmer Genomics.