821

The Role of Whole Genome Sequencing in Clinical Diagnosis

William A. Langley, PhD, Nicholas A. Rouse, MS, Karan A. Kapuria, MS, Rebecca Kelly, BSc, Daniele Alaimo, MSc, Ymkje Cuperus, MS, James Chappell, PhD, Peter L. Nagy, MD, PhD
MNG Laboratories, Atlanta GA

MNG LABORATORIES
Neurogenetic Answers™

I. Introduction

Whole exome sequencing (WES) is often regarded as the clinical gold standard for diagnostic testing in cases where a genetic etiology is suspected. This methodology allows for the sequencing of all the exons, or protein-coding regions, of an individual's DNA. However, these exonic regions account for only 1.5% of an individual's entire DNA sequence. As scientific knowledge has advanced, it has become evident that variants in non-coding regions can also play an important role in genetic disorders. Whole genome sequencing (WGS) allows for the detection of variants within these non-coding regions. To assess the functional consequences of these variants, transcriptome analysis can be used to improve the sensitivity and specificity of WGS.

Using the MNGenome® test recently developed in our laboratory, we have identified multiple cases in which WGS with simultaneous transcriptome analysis resulted in a likely clinical diagnosis for patients for whom WES would have been non-diagnostic. In addition to single nucleotide variants and large copy number variants (CNVs), this test can also detect smaller indels/CNVs in the 250bp-50Kb range, mitochondrial depletion, and targeted repeat expansions. To further investigate these findings, we performed MNGenome® plus transcriptome analysis testing on 22 cases in which WES testing was previously non-diagnostic, but a genetic etiology was highly suspected.

Our findings lead us to the conclusion that whole genome sequencing, coupled with transcriptome analysis, is a superior genetic testing modality and offers a diagnosis in multiple cases where other methods fail to provide diagnostic or prognostic insight into a patient's condition. By combining WGS with transcriptome analysis, we have created a new gold standard for comprehensive diagnostic testing.

II. Methods

MNGenome® Analysis (Dragen; PMID: 26419432)
- Mitochondrial genome sequencing, deletions, heteroplasmy and depletion
- Greater than 99.4% of uniquely mappable regions covered 30-fold
- Variants called in RNA genes
- All pathogenic variants covered
- CNVs across the genome
- CNVs with 250-bp resolution
- Repeat expansion detection using Expansion Hunter (Dolzhenko et al. 2017)

Transcriptome Analysis
- Illumina TruSeq strand-specific library prep
- Analysis by HISAT2, StringTie: splicing-aware alignment, reconstruction, and quantification of transcripts (PMID: 25751142)
- Differential Expression: Compare gene expression versus controls
- Transcript Junction analysis: Determine junctions that are missing or novel, identification of novel transcripts
- Variant calling from transcriptome data using RNA specific software

Integrative Analysis
- Search for genomic variants (SNV & CNV) in coding and regulatory regions
- Transcriptome helps interpret the significance of genetic changes by revealing what is actually expressed in a tissue
  - •Regulatory mutations - enhancer and promoter mutations
  - •Large deletions of coding region --transcription is decreased by half, at imprinted loci transcript is absent
  - •Intragenic deletions -Alternative initiation and termination sites, exon skipping and nonsense mediated decay
  - •Identification of artifacts due to pseudogenes in exome data

III. Results

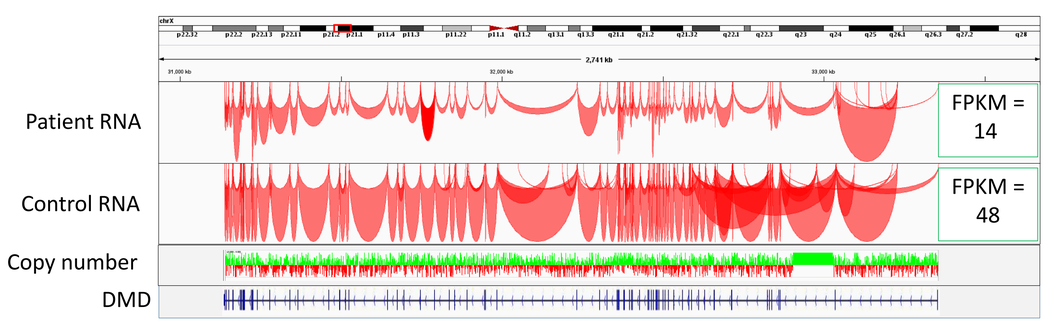

Figure 1: A 9-year-old male presenting with muscle weakness, increased serum creatine kinase and abnormal immunohistochemistry of muscle with increased myofiber size variation and atrophic/hypertrophic fibers. Becker muscular dystrophy was the suspected diagnosis; however previous sequencing and del/dup analysis of the DMD gene was negative. Using the MNGenome® we identified a duplication in the second intron. Integration of transcriptome testing identified the functional consequence, whereby transcript expression levels are decreased in the patient sample compared to the control likely by affecting RNA processing or stability.

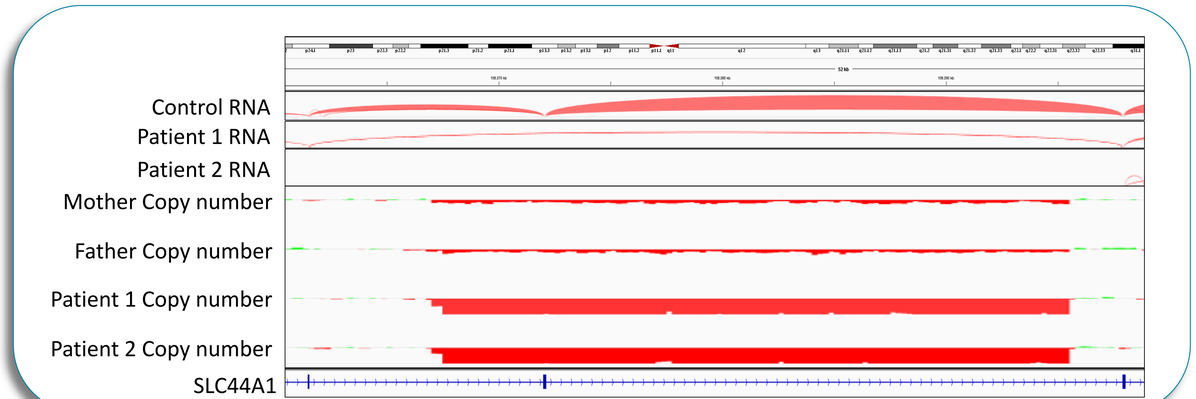

Figure 2: Two similarly affected siblings. An 18-year-old male presenting with progressive motor weakness, developmental regression, abnormal MRI and a thin, elongated face and his 22-year old sister presenting with decreased visual acuity, ataxia, dysphagia, progressive neurological disorder and progressive change in face shape. Using the MNGenome® we identified a copy number deletion in the third exon of SLC44A1. Integration of transcriptome testing identified the functional consequence, whereby exon 3 is skipped during transcription, which is predicted to cause a downstream frameshift and premature termination during translation.

IV. Conclusion

Combined genome/transcriptome testing helps interpret:
- Regulatory mutations - enhancer and promoter mutations
- Large deletions of coding regions - transcription is decreased by half and at imprinted loci transcript is absent
- Intragenic deletions - alternative initiation and termination sites, exon skipping and nonsense mediated decay or lack thereof
- Identification of artifacts due to pseudogenes
- Processing outcome of specific variants to assist in the development of splicing modifiers

Of the 22 nondiagnostic exome trios selected, we identified a likely clinical diagnosis in 4 cases.
WGS coupled with transcriptome analysis is a superior genetic testing modality and offers a diagnosis in a significant number of cases where other methods fail to provide diagnostic or prognostic insight into a patient's condition.

LabCorp

V. References

1. A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. Miller et al, Genome Med 2015 (PMID: 26419432)
2. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Pertea et al, Nat Biotechnol 2015 (PMID: 25690850)
3. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Pertea et al, Nat Protoc 2016 (PMID: 27560171)
4. Diagnostic Yield and Novel Candidate Genes by Exome Sequencing in 152 Consanguineous Families With Neurodevelopmental Disorders. Reuters et al, JAMA 2017 (PMID: 28097321)