MNG LABORATORIES
Neurogenetic Answers™

RNA Sequencing: Improving Diagnostic and Therapeutic Outcomes for Genetic Testing

W.A. Langley, Y. Cuperus, N.A. Rouse, Jiyong Wang, P.L. Nagy

Introduction

- Next generation sequencing of DNA is extremely powerful in identifying the cause of genetic disorders. Finding a rare, "in silico" disruptive variant is not sufficient to establish its pathogenicity. Reporting of variants of unknown significance (VUS) causes anxiety and increases the workload of the care providers. Functional tests, such as transcriptome analysis/ RNA sequencing can help in the evaluation of the molecular consequences of both SNVs and CNV.
- RNA-sequencing[1] can demonstrate the pathogenic nature of deletions, frameshifts or non-canonical splice-site variants by revealing the effects of these changes on the level and content of transcripts present in the affected tissue. It is even useful in the diagnoses of imprinting disorders, irrespective of the underlying molecular mechanism. If the gene is widely expressed, blood or other easy to sample tissue types can be used with similar efficiency. Transcriptome analysis also helps to identify patients who could benefit from disease therapies based on splice-modulating therapeutic approaches.
- We have developed validated and have been using clinical transcriptome sequencing to assist with the diagnosis of constitutional disorders since last November. In our experience transcriptome analysis can rule in or rule out the proposed splicing disruptive effect of VUS's in almost all cases where appropriate tissue is available. In addition, in some cases it can also provide diagnosis in the absence or instead of DNA-based diagnostic approaches (Prader Willi syndrome and SMA testing).

Why Perform Transcriptome Testing - Summary

- Transcriptome is the summary of all different genetic changes
  - Regulatory mutations
    - Enhancer and promoter mutations
  - Large deletions of coding regions
    - Transcript level is decreased by half
    - In imprinted loci transcript is absent
  - Intragenic indels
    - Alternative initiation and termination sites
    - Exon skipping
    - Nonsense mediated decay
  - Intragenic SNVs
    - Splice site efficiency modulation
    - Premature stops activating nonsense mediated decay
- With the development of splicing modifiers, it is important to assess the exact processing outcome of a specific variant

Methodology - Design

- Next generation sequencing library preparation by TruSeq Stranded Total RNA Library Prep Kit/ Ribo-Zero Human (Illumina)
- Validated for blood, muscle, brain/nerve, fibroblast culture
- RNA sequencing analysis is run on a HiSeq or NovaSeq instruments (Illumina)
- Housekeeping genes were determined as the set of genes that had average expression and low variance across all tissue types
- Reportable genes were defined as genes that have expression > FPKM1 and low variance in a given tissue
- Reference ranges were established for junction, gene, and transcript expression using the tissue specific control sets
- To determine the preferred sample type, the human protein atlas[3] is used as one of the sources to look at gene expression in different tissues

Quality control of RNAseq samples

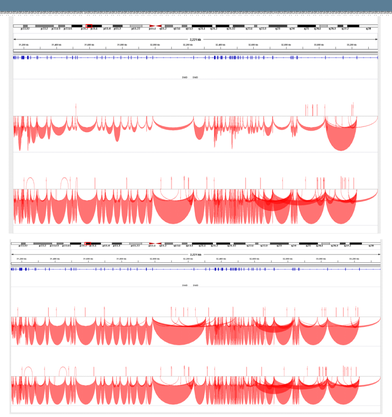
- High quality reads: > 85 % of all reads Q30 or greater
- Total yield per sample: > 10 Gb
- Alignment statistics
  - Alignment rate: > 90 %
  - Number of uniquely mapped reads: > 50 million
  - rRNA contamination percentage: < 2 % of all mapped reads
- Other metrics tracked
  - Transcript reconstruction statistics
  - Exon-exon junction reconstruction statistics

Bioinformatics

- Tools used:
  - HISAT (hierarchical indexing for spliced alignment of transcripts), StringTie and Ballgown http://ccb.jhu.edu/software.shtml.
    - Aligns reads to a genome
    - Assembles transcripts including novel splice variants
    - Compute the abundance of these transcripts and splice junctions in each sample
    - Compare experiments to identify differentially expressed genes and transcripts

Visualization with Integrated Genome Viewer (IGV)

Fig 1: DMD diagnosis



I: 9 year old male patient with muscle weakness
Serum CK ~13,000
Immunohistochemistry of muscle: Increased myofiber size variation, broad range of atrophic and hypertrophic fibers
Negative deletion/duplication and sequencing analysis
Family history: maternal uncle diagnosed with Becker Muscular Dystrophy
Sample type: muscle, gene DMD

Results: Positive, differential expression of DMD
Total transcript level change per gene: -13.1 FPKM (DMD)
Transcript log2 fold change per gene: -2.2 (DMD)
Splice site usage changes/exon skipping: Negative

II: 3 year old male patient with muscle weakness
Immunohistochemistry of muscle: Absence of Dystrophin
Deletion/duplication analysis of DMD: deletion of exon 44
Family history: negative
Sample type: muscle, gene DMD

Results: Positive, Deletion of exon 44
Total transcript level change per gene: within reference range
Transcript log2 fold change per gene: within reference range
Splice site usage changes/exon skipping: Skipping of exon 44 (chrX32235033-3225180; Transcript ID: NM000109; Amino Acid coding number: 2097-2146)

Prader Willi Syndrome

- Prader Willi critical region on 15q11.2-q13
- Abnormal parent-specific imprinting, 3 known mechanisms: paternal deletion, uniparental disomy 15 (UPD), and imprinting defect (ID)
- Severe hypotonia and feeding difficulties early in life
- Later excessive eating and central obesity
- Delay in motor milestones and language development
- Cognitive impairment, with behavior problems
- Short stature
- Hypogonadism, incomplete pubertal dev, and infertility

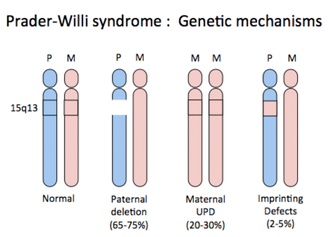
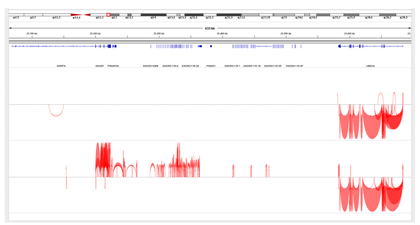


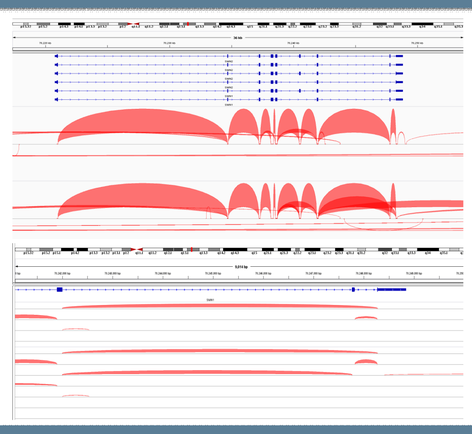
Fig 2: Prader- Willi diagnosis



I: 17 year old male patient
Clinical symptoms suggestive of Prader Willi Syndrome
Family history: negative
Previous testing: altered methylation and loss of paternal expression at 15q
Sample type: whole blood, gene SNRPN

Results: Positive, no expression
Total transcript level per gene: -20.15 FPKM (mean: 20.15; std: 6.29) (SNRPN)
Transcript log2 fold change per gene: NA (no expression)
Splice site usage changes/exon skipping: Negative

Fig 3: SMA diagnosis



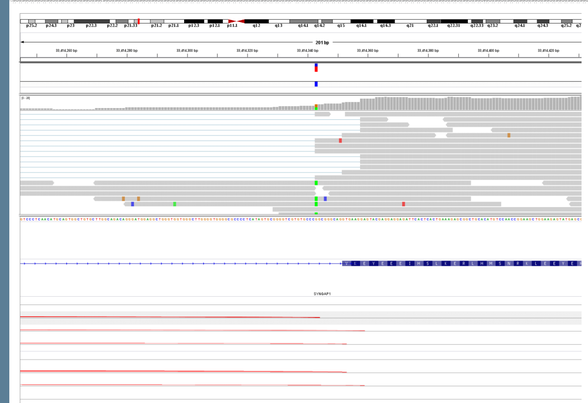
I: 2 year old female patient
Clinical symptoms consistent with SMA
Family history: negative
Sample type: blood, gene SMN1/SMN2

Results: Positive,
Splicing of SMN2

II: 1 year old female patient with a clinical diagnosis of spinal muscular atrophy
Family history: negative
Sample type: whole blood, gene SMN1/SMN2

Results: Positive for a deletion of exon 7 in SMN1

Fig 4: Assessment of splice sites VUS



I: 2 year old male patient
Generalized epilepsy
Confirmation requested of a SYNGAP1 variant (c.3583-9G>A; Chr6:33414343)
Classified as pathogenic
Family history: negative

Sample type: whole blood

Results: Positive

Presence of variant confirmed in 5 of the 6 hnRNA reads.
In addition, 5 incorrectly spliced transcripts between exons 16-17 (Chr6:33412394-33414345) resulting in a frameshift and premature termination of protein synthesis

Pathogenicity of variant confirmed

RNA Sequencing Test Options

MNG Transcriptome™	Complete RNA transcriptome sequencing that can be ordered at the time of initial testing as a reflex to any NGS panel, or in conjunction with our MNG Exome™.
Panel-Specific RNA Sequencing	RNA sequencing analysis can be ordered as an add-on to any of MNG Laboratories' NGS panels and will only include analysis of the genes listed in the panel.
Gene-Specific RNA Sequencing	Analysis of up to five genes for variant investigation following a report with a VUS identified in a splice site or intronic region, or the variant is noted to have a potential effect on splicing. Previous sequencing analysis can be done at any laboratory.

Determining Tissue Type for Testing

MNG uses multiple criteria and a team approach to help determine the best sample type to submit based on clinical relevance to the patient and variants

Internally developed RNA expression profiles and reportable ranges across all genes and tissue types	Publicly available databases for protein and RNA expression	Clinical information relevant to the patient's phenotype and symptomology

Tissue Specific Shipping Conditions

Tissue Types	Blood	Fibroblasts and Lymphoid Cell Lines	Nerve, Brain, Muscle Tissue
Sample Preparation and Collection	MNG Provided Tempus™ Blood RNA Collection Tube (Thermo Scientific)	Cultured, T-25 flask Preferred: cell pellet spun down and frozen	50-75 milligrams of fresh, frozen tissue
Shipping Conditions	Room temperature, overnight shipping	5-7 pounds dry ice, overnight shipping	5-7 pounds dry ice, overnight shipping

Please call MNG Labs at 678-225-0222 for any questions or assistance

References

1. Cummings et al, Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med, 2017; 9(386) PMID 28424332
2. Pertea et al, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc, 2016; 11(9): 1650-1667 PMID 27560171
3. www.proteinatlas.org

MNG Laboratories I 5424 Glenridge Drive NE, Atlanta, GA 30342 USA I +1 678.225.0222 I mnglabs.com