

James Chappell¹, Michael Reott¹, Rebecca Kelly¹, Daniele Alaimo¹, Ymjke Cuperus¹, Trey Langley¹, Peter Nagy¹

1. MNG Laboratories, Laboratory Corporation of America®

I. Introduction

Whole exome sequencing is a standard approach for identifying genetic variants underlying neurological disorders. Despite its proven utility, whole exome testing often identifies variants of uncertain significance, resulting in an indeterminate test result. We have found that including RNA sequencing analysis as a complementary method has been useful for identifying the functional consequences of variants of uncertain significance and has improved our diagnostic rate. We describe here how inclusion of whole transcriptome sequencing with DNA sequencing for multiple patients with neurological disorders facilitated molecular diagnoses.

II. Case Descriptions

Case 1 (Figure 1) is a 36-year-old patient with a history of intellectual disability, optic atrophy, and ataxia. Previous MRI findings were suggestive of Leigh disease, and additional testing was consistent with a mitochondrial complex deficiency. DNA sequencing revealed a homozygous splice site variant in the ACAD9 gene, c.244+3A>G (Chr3:128603592, GRCh37). This variant was determined to be heterozygous in the patient's mother and father. The ACAD9 gene is associated with autosomal recessive mitochondrial complex I deficiency, nuclear type 20 (OMIM:611126).

Cases 2 and 3 (Figure 2) are a pair of siblings. Sibling one is an 18-year-old patient presenting with progressive motor weakness, proximal muscle weakness, positive Gower's sign, developmental regression, MRI abnormalities, and a thin, elongated face. Sibling two is 22-year-old patient presenting with decreased visual acuity, ataxia, tremor, dysphagia, cerebellar volume loss, and changes in facial shape and movement. DNA sequencing for both siblings revealed a homozygous deletion involving exon 3 of the SLC44A1 gene (Chr9:108066751-108095501, GRCh37). This variant was determined to be heterozygous in the patients' mother and father. The SLC44A1 gene is reported to be involved in neurogenesis, and a recent publication describes a loss-of-function variant in a patient with a similar neurological phenotype (PMID:28097321).

Case 4 (Figure 3) is a 12-year-old patient presenting with tuberous sclerosis. DNA sequencing revealed a splice site variant in the TSC2 gene, c.1599+5G>A (Chr16:2114433, GRCh37).

Case 5 (Figure 4) is a 5-year-old patient with generalized epilepsy. DNA sequencing revealed a splice site variant in the SYNGAP1 gene, c.3583-9G>A (Chr6:33414343, GRCh37).

III. Methods

MNG Exome™ Analysis (Dragen; PMID: 26419432)

- Agilent SureSelect Human Exome V6 capture
- Individual capture of samples allows CNV detection
 - >200-fold average coverage
 - >97% ROI covered 30-fold
 - Allows variant calling on >99% ROI
- All pathogenic variants at or above "expert consensus" level (ClinVar) covered

Transcriptome Analysis

- Illumina TruSeq strand-specific 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 missing/novel junctions, novel transcripts
- Variant calling from transcriptome data using GATK (version 4.0.12.0)

Integrative Analysis

- Search for 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

IV. Results

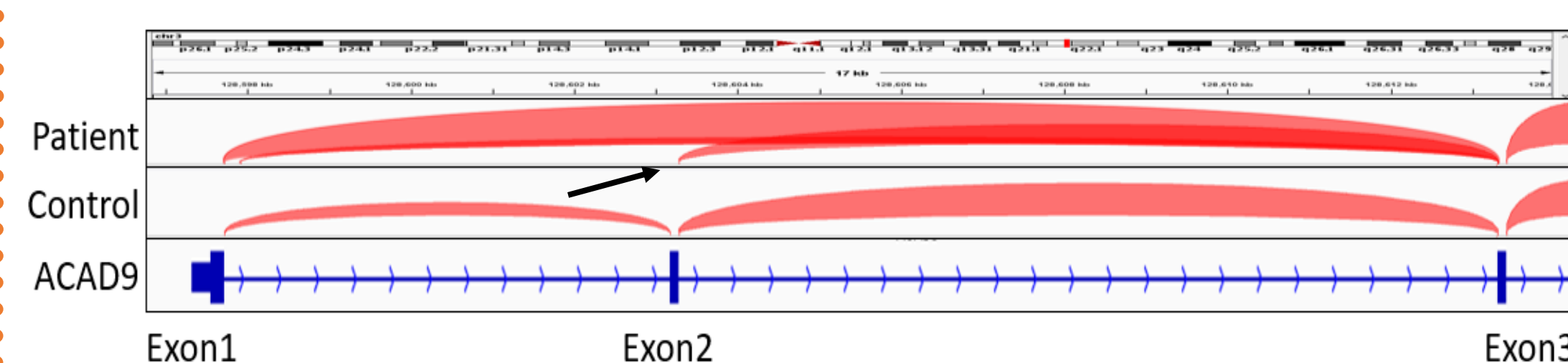


Figure 1: Functional Consequences of a Novel Splice Site Variant in ACAD9. A homozygous splice site variant was detected in ACAD9 in exon 2/18 at c.244+3A>G. This variant is the change of the adenine at the +3 nucleotide within the donor splice region of exon 2 to a guanine. RNA studies indicate that the variant results in skipping of exon 2, which results in a frameshift at amino acid position 82. A termination codon is predicted 20 amino acids beyond this change. The data also indicate low levels of splicing from exon 2 to exon 3, but this transcript is lacking exon 1, and its functionality is unclear. Potentially, it could result in a N-terminal truncated protein with some residual functionality. To our knowledge, this variant has not been reported previously.

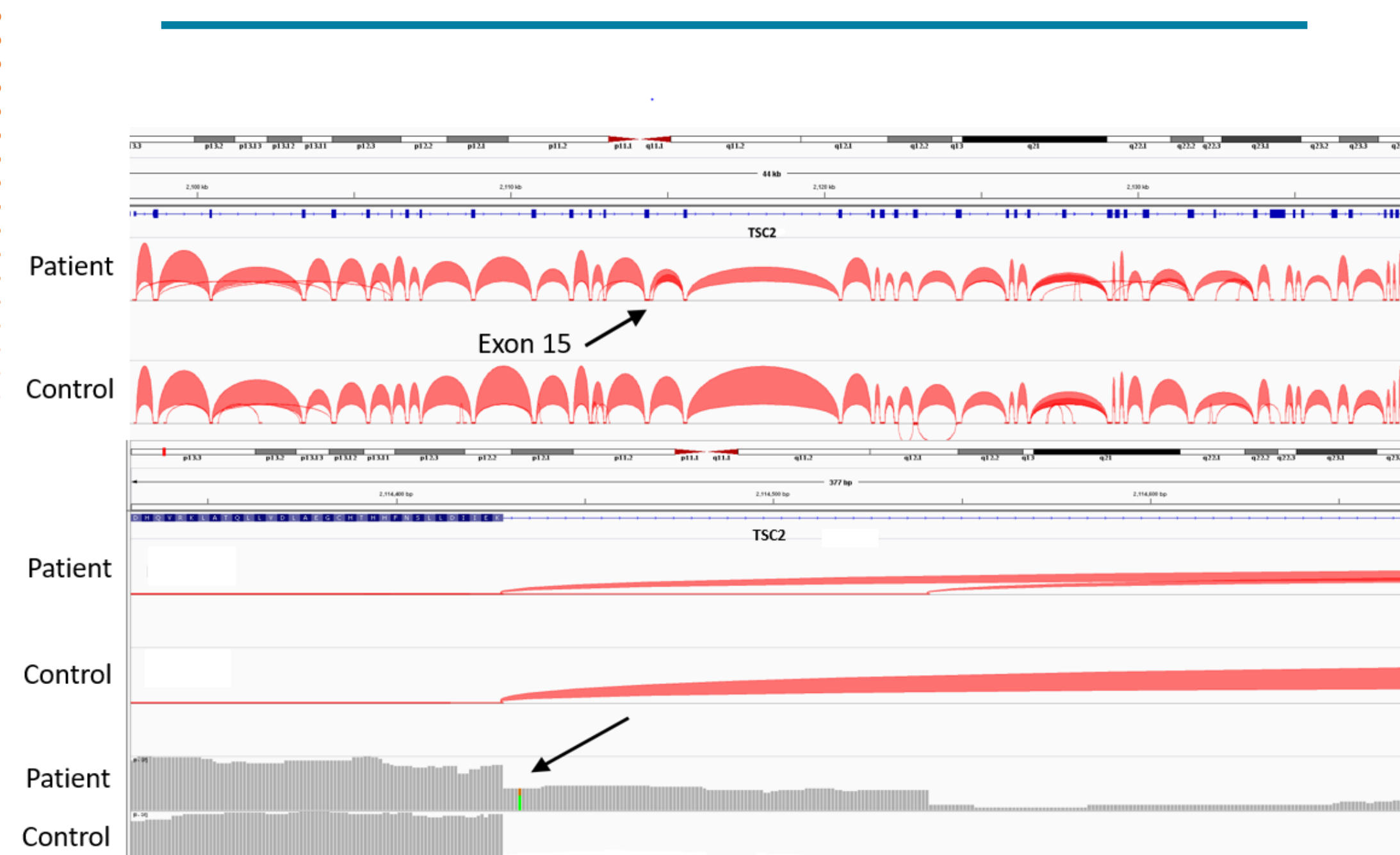


Figure 3: Functional Consequences of a Reported Splice Site Variant in TSC2. A splice site variant was detected in TSC2 in exon 15/42 at c.1599+5G>A. This variant is the change of the guanine at the +5 nucleotide within the donor splice region of exon 15 to an adenine. This variant has previously been reported; however no definitive classification has been assigned. RNA studies indicate that the variant creates a novel splice junction from intron 15 to exon 16 (Chr16:2114542-2115520). This novel splice site is predicted to result in the insertion of 10 amino acids followed by a stop codon, resulting in a truncated protein.

V. Conclusion

DNA sequencing often provides the answer to diagnostic dilemmas, but in some situations, functional assessment of variants is necessary that can be provided by transcriptome sequencing. We investigated the use of RNA-seq analysis as a complementary method to investigate interesting cases which had indeterminate molecular diagnoses by DNA sequencing alone. In the first case, a potential splice site variant was determined to have functional consequences once we visualized exon-skipping in the RNA. In the second case, a pair of siblings with a deletion detected by exome sequencing were investigated. In each of their RNA analysis, it was determined that the exon was missing from the transcripts RNA sequencing. In the third case, a +5 splice site variant was determined to create a novel splice site. Finally, in the fourth case, a -9 splice site variant that had previously been described in patients with epilepsy was determined to have functional significance. Without additional functional assessment of these variants, the cases described would have left both the clinician and the patient with an indeterminate report, highlighting the clinical utility of incorporating transcriptome sequencing as a complementary method.

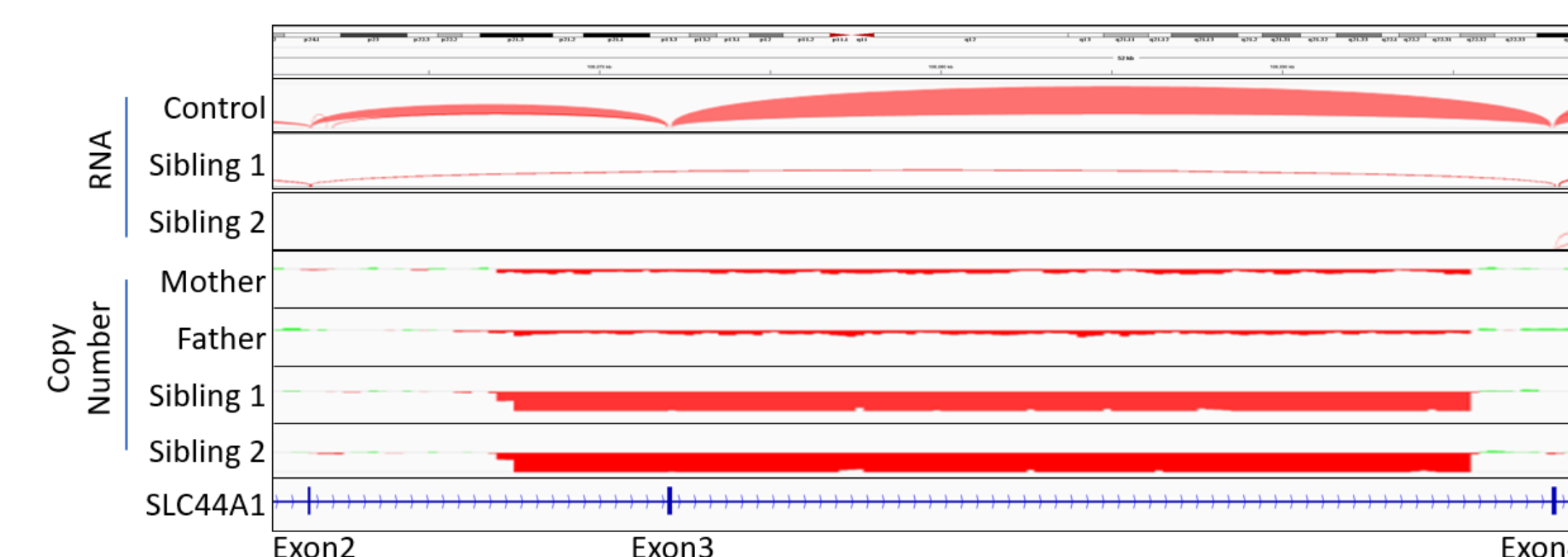


Figure 2: Functional Consequences of a Novel Deletion in SLC44A1. The figure shows a zoomed in view of exons 2, 3, and 4 and the associated introns for SLC44A1. Copy number analysis identified a homozygous deletion which includes exon 3 in both patients. The functional consequences of this deletion are observed in the patients' RNA where they exhibit skipping of exon 3. This is predicted to cause a frameshift and downstream termination event, and decreased expression. To our knowledge, this variant has not been reported previously.



Figure 4: Functional Consequences of a Reported Splice Site Variant in SYNGAP1. A splice site variant was detected in SYNGAP1 in exon 17/19 at c.3583-9G>A. This variant is the change of the guanine at the -9 nucleotide within the acceptor region of exon 17 to an adenine. RNA studies indicate that the variant disrupts normal splicing. The novel splice site is predicted to result in a change from a valine to an arginine at amino acid position 1195 and a frameshift thereafter. A termination codon is predicted 27 amino acids after this change, resulting in a truncated protein. This variant has been reported in patients with epilepsy; however the functional significance of this variant was previously unknown.

VI. 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. HISAT: a fast spliced aligner with low memory requirements. Kim et al, Nat Methods 2015 (PMID: 25751142)
3. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Pertea et al, Nat Biotechnol 2015 (PMID: 25690850)
4. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Pertea et al, Nat Protoc 2016 (PMID: 27560171)
5. Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. Reuter et al, JAMA Psychiatry 2017 (PMID: 28097321).