Child Neurology: RNA Sequencing for the Diagnosis of Lissencephaly

Hebah Qashqari1, Arun Ramani5, Hernan Gonorazky1,2, Kimberly Amburgey1,2, Mohammad M. Ghahramani Seno2, Michael Brudno2,5,6,7, Sergey Naumenko8, Soma Das9, James J. Dowling1,2,3,4

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Neurology® Published Ahead of Print articles have been peer reviewed and accepted for publication. This manuscript will be published in its final form after copyediting, page composition, and review of proofs. Errors that could affect the content may be corrected during these processes.
1. Division of Neurology, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Ontario, Canada
2. Program of Genetic and Genome Biology, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
3. Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada
4. Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
5. Centre for Computational Medicine, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
6. Research Department of Computer Science University of Toronto, Ontario, Canada
7. Techna Institute, University Health Network, Toronto, Canada
8. Department of Biostatistics, Harvard Chan School of Public Health, Boston, MA, USA
9. Department of Human Genetics, University of Chicago, Chicago, IL, USA

CORRESPONDING AUTHOR:
James Dowling
E-mail: James.Dowling@sickkids.ca

WORD COUNT: 748

FIGURES: 1
REFERENCES: 7

KEYWORDS:
Lissencephaly, PAFAH1B1-Associated Lissencephaly, Platelet-activating factor acetylhydrolase Isoform 1B, Alpha-subunit (PAFAH1B1), LIS1-associated Lissencephaly, RNA sequencing.

CONSENT: An informed consent was obtained from the parents for the publication of this case report and accompanying images.

AKNOWLEDGEMENTS: The authors thank the patient and his family for their participation.

STUDY FUNDING: GenomeCanada DIG phase II (MB, JJD)

DISCLOSURE: H. Qashqari reports no disclosures relevant to the manuscript. A. Ramani reports no disclosures relevant to the manuscript. H. Gonorazky reports no disclosures relevant to the manuscript. K. Amburgey is an employee at Deep Genomics. M. M. G. reports no disclosures relevant to the manuscript. M. Brudno is an employee at Gene42 Inc. company. S. Naumenko reports no disclosures relevant to the manuscript. S. Das reports no disclosures relevant to the manuscript. J.J Dowling reports no disclosures relevant to the manuscript.
Introduction

Malformations of cortical development represent an important cause of developmental disability and neurologic morbidity and mortality(1). Advances in genetic methodology, particularly the widespread implementation of “next generation” DNA sequencing technology (e.g. multi gene panels and whole exome sequencing), have significantly improved diagnostic yield in neurogenetic disease(2). The current yield for a range of conditions, including brain malformations, epilepsy(3), global developmental delay, and movement disorders, is approximately 50%.

Despite these advances, there are still many individuals with suspected genetic disease who remain genetically unresolved. There are likely multiple explanations for this, with one important contributor being pathogenic variants(s) in the non-coding genome that impact RNA processing and transcript expression. Whole transcriptome analysis, as accomplished using RNA sequencing (or RNAseq), is an emerging technology suitable for uncovering RNA processing mutations(4), as has been successfully demonstrated in particular for neuromuscular diseases(5, 6).

In this study, we use RNAseq to aid in the diagnosis of an individual with classical (type I) lissencephaly and a reportedly normal brain malformation multi gene panel. Transcriptome analysis uncovered a splice altering variant, subsequently verified by DNA sequencing, in the LIS1 gene. We conclude that RNAseq represents an effective diagnostic tool for genetically undefined cases of cortical malformation that can greatly improve the current diagnostic rate.

Case report:

We present a 3-year-old boy born at term to non-consanguineous Portuguese parents after an uneventful pregnancy and delivery. First medical concerns were at age 6 months when he presented with infantile spasms (epileptic spasms plus EEG with hypsarrhythmia). Currently, he has developmental delay with superimposed developmental regression in the context of medically refractory epilepsy. His physical examination is notable for microcephaly 45.5cm (-2SD), dysmorphisms including hypertelorism and a depressed nasal bridge, axial hypotonia and exaggerated deep tendon reflexes. Brain MRI (performed at six months) is consistent with classical (type I) lissencephaly with a posterior-to-anterior gradient (Figure 1A).

On the basis of these features, there was high clinical suspicion for an underlying genetic condition, and specifically a defect in the LIS1 (or PAFAH1B1) gene(7). Commercial genetic testing was performed, and included an epilepsy multi gene panel (471 genes including LIS1, Courtagen) and MLPA (multiplex ligation dependent probe amplification)-based analysis for deletion/duplication of LIS1 (University of Chicago). No causative abnormality was discovered through these approaches.

Based on the negative DNA testing, we pursued RNA sequencing through an REB approved research protocol. As source material is a critical consideration for RNAseq, and since no brain tissue was available in this case, we wanted to understand which easily accessed tissues best mirrored the brain transcriptome. We focused on blood and skin fibroblasts, using transcriptome data from our “in house” database and from GTex. While neither had profiles that fully matched the brain, fibroblasts much more closely resembled cortex in terms of number of genes expressed and expression levels (Figure 1C). In particular, we reviewed tissue expression of LIS1 and other known brain malformation genes, and found they have substantially higher expression in fibroblasts, with levels nearly equivalent to cortical brain tissue.
We thus performed a skin biopsy, derived fibroblast cultures, and performed RNAseq. The resulting total transcriptome analysis revealed reduced expression and allele imbalance of LIS1, with approximately 20 reads coming from one allele versus > 500 from the other. Within the allele with reduced expression, we observed multiple splicing disruptions, including exon 4 skipping, though the low overall read count presented difficulty for the definitive interpretation of specific splicing alterations. Overall, the RNAseq data was consistent with reduced expression of one LIS1 allele, supporting a haploinsufficiency mechanism, the most common mutational consequence associated with LIS1 mutation.

We repeated DNA sequence analysis with Sanger based sequencing of LIS1 gene on a clinical basis. This demonstrated a pathogenic variant in exon 4 of LIS1 gene with a sequence change at c.164 G>A. This variant is predicted to create a premature stop codon at amino acid residue 55, p.Trp55*, with subsequent nonsense mediated decay. Our transcriptome data supports the reduction of expression from this allele, and suggests the variant promotes splicing changes in addition to nonsense mediated decay. Given that heterozygous loss of expression/function variants in LIS1 are the known genetic pattern for LIS1 related lissencephaly, we thus concluded that this heterozygous stop variant is the cause of disease in our patient.

Discussion:

Our study demonstrates the suitability of RNAseq for identifying novel mutations, and also for providing functional evidence of mutation consequence. Using RNAseq, we were able to accurately identify a variant in LIS1, and show that the resulting nonsense mutation promoted allele specific loss of expression and abnormal RNA processing. Importantly, while our RNAseq analysis ultimately pointed to a pathogenic variant in LIS1, because we captured the total transcriptome, we were potentially able to identify an abnormality in any gene expressed in the sample. For fibroblast transcriptome analysis, this includes essentially all genes associated with cortical malformations, as well as most genes associated with neurogenetic conditions. Overall, therefore, this technology is applicable for the investigation of any case not fully clarified or solved by multi-gene panel or whole exome sequencing.

In terms of RNA sequencing as a diagnostic modality, our study reinforces the importance of using suitable source material for transcriptome analysis, and provides first proof of concept for using skin fibroblasts for the study of brain malformation genetics. This is particularly useful for clinical practice, given that brain tissue is often not available, and skin biopsies are an efficient and non-invasive bedside procedure that can assist in revealing an underlying genetic diagnosis. Of note, we show that the expression of genes associated with brain malformations is superior in fibroblasts as compared to blood, both in terms of number of genes expressed and overall levels of expression. This is important because only genes expressed at suitable (typically >1 transcript per million (TPM) for expression and at least 5 reads for identifying splicing aberrations) can be meaningfully analyzed. Another potential source material for diseases of the central nervous system are lymphoblastoid cell lines, which have recently been shown to express a broad range of genes associated with neurodevelopmental disorders, and which have been successfully used to identify and/or clarify pathogenic variants that cause Cornelia de Lange Syndrome(8).

Additionally, this case points to an important limitation of some genetic testing platforms, as the LIS1 variant in this case was missed in the initial analysis of patient DNA. The reason(s) why the variant was not detected are not clear, as overall mean read depth of the gene in the panel was >
500x. One possibility is that the specific coverage of exon 4 was low, a likely hypothesis given that exon 4 is known to be challenging to capture at sufficient depth by next generation sequencing. Overall, this illustrates the critical need to understand strengths and limitations of different genetic diagnostic technologies, and to not fully exclude causes until a definitive diagnosis has been established.

A final consideration is when to utilize RNAseq in the diagnostic pathway. Multi-gene panel or whole exome sequencing remain first line testing methodologies for mutation identification across the spectrum of neurogenetic disorders. However, these technologies do not provide information related to the non-coding genome, and additionally often identify variants that are of uncertain significance. RNAseq is therefore an ideal “next step” modality in panel or WES negative cases and/or cases with VOUSs, as it provides data on variants (coding or non-coding) that impact RNA levels and processing. Moving forward, with the emergence of whole genome sequencing (WGS) as a powerful and comprehensive technology for detection of pathogenic variants(9), RNAseq is also well positioned to provide functional annotation and interpretation when paired with WGS.

Methods:
RNA sequencing
Fibroblasts were derived from dorsal palmar skin biopsy, expanded through at least 2-3 passages, and then prepared for RNA extraction. RNA was extracted using Qiagen RNAeasy mini kit. Libraries for RNAseq were prepared using poly-A selection (Illumina TruSeq) at The Centre for Applied Genomics (TCAG, SickKids), and paired-end 126+126 bp sequencing was subsequently performed with Illumina HiSeq 2000 instruments at TCAG. RNAseq data generation and analysis were carried out as described (6).

Sequence alignment, expression, variant calling and splicing
Alignment, variant calling, and quality controls steps were carried out using the RNA-seq workflow from the bcbio-nextgen bioinformatics framework (version 1.1.0). Raw reads were aligned to the GRCh37 (hg19) version of the human reference genome using the splice-aware aligner STAR. Expression was calculated using the R-bioconductor package edgeR and variants were called using GATK best-practices. Finally, novel and outlier splice junctions were identified using the rules and filters outlined previously (6). The resulting BAM was viewed using Integrative Genomics Viewer to generate sashimi plots and identify aberrant splicing changes.

Heat map generation
We used the GTEx “multi gene query” portal (GTEx Portal) to generate a gene expression heatmap. For the 42 brain malformation genes (listed in Figure 1C, identified from the GeneDx panel), we selected tissues of interest (all brain tissues, fibroblasts, whole-blood, skeletal muscle, and skin) and exported a svg (scalable vector graphics) of the resulting heatmap. The image was used to generate Figure 1C where we highlight LIS1 (the gene of interest).
References:
Figure 1. 3-year-old boy with posterior-anterior gradient lissencephaly.
A) Sagittal T1 MRI showing Dobyns grade 3 lissencephaly more prominent in the parietooccipital region with posterior-anterior gradient and normal midline structure. B) Axial T2 MRI at the level of the thalamus showing agyria/pachygyria in the parietooccipital region. C) Heat map coverage of common brain malformation genes in fibroblast vs. blood vs. brain tissue. Expression (in Transcripts per Million – TPM) of 42 brain malformation genes in each tissue is represented as a heatmap with colors ranging from yellow (0 TPM) to black (6.7 e+3 TPM). Most of the 42 genes are expressed at higher levels in fibroblasts (column 3) compared to whole blood (column 1). Additionally, the expression in fibroblast is more consistent with expression in brain tissues. D) Sashimi plot showing transcript change with exon 4 skipping. We show expression of exons 3, 4 and 5 of LIS1 in the proband (case) and two unrelated controls. The coverage across each exon is plotted as a bar graphs, arcs represent splice junctions connecting exons and the number in the arc shows the number of reads split across the junction. The arrow points to the 20 reads connecting exons 3 and 5 (skipping exon 4) in the proband that are not found in either of the control samples.
Child Neurology: RNA Sequencing for the Diagnosis of Lissencephaly
Hebah Qashqari, Arun Ramani, Hernan Gonorazky, et al.
Neurology published online May 20, 2021
DOI 10.1212/WNL.0000000000012265

This information is current as of May 20, 2021

Updated Information & Services
including high resolution figures, can be found at:
http://n.neurology.org/content/early/2021/05/20/WNL.0000000000012265.citation.full

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.neurology.org/about/about_the_journal#permissions

Reprints
Information about ordering reprints can be found online:
http://n.neurology.org/subscribers/advertise