

Child Neurology: RNA Sequencing for the Diagnosis of Lissencephaly

Hebah Qashqari, MBBS, FRCPC, Arun Ramani, PhD, Hernan Gonorazky, MD, Kimberly Amburgey, MSc, Mohammad M. Ghahramani Seno, DVM, PhD, Michael Brudno, PhD, Sergey Naumenko, PhD, Soma Das, PhD, and James J. Dowling, MD, PhD

Correspondence

Dr. Dowling
James.Dowling@sickkids.ca

Neurology® 2021;97:e1253-e1256. doi:10.1212/WNL.00000000000012265

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., multigene panels and whole exome sequencing [WES]), 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 many individuals with suspected genetic disease who remain genetically unresolved. There are likely multiple explanations for this, with one important contributor being pathogenic variants in the noncoding genome that affect RNA processing and transcript expression. Whole transcriptome analysis, as accomplished using RNA sequencing (RNA-seq), 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 used RNA-seq to aid in the diagnosis of an individual with classical (type I) lissencephaly and a reportedly normal brain malformation multigene panel. Transcriptome analysis uncovered a splice altering variant, subsequently verified by DNA sequencing, in the *LIS1* gene. We conclude that RNA-seq represents an effective diagnostic tool for genetically undefined cases of cortical malformation that can greatly improve the current diagnostic rate.

Case Report

A 3-year-old boy was born at term to nonconsanguineous 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 was notable for microcephaly 45.5 cm (−2 SD), dysmorphisms including hypertelorism and a depressed nasal bridge, axial hypotonia, and exaggerated deep tendon reflexes. Brain MRI (performed at 6 months) is consistent with classical (type I) lissencephaly with a posterior-to-anterior gradient (Figure, A).

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

From the Division of Neurology, Department of Pediatrics (H.Q., H.G., K.A., J.J.D.), Centre for Computational Medicine (A.R., M.B.), and Program of Genetic and Genome Biology (H.G., K.A., M.M.G.S., M.B., J.J.D.), the Hospital for Sick Children, and Department of Pediatrics (J.J.D.), Department of Molecular Genetics (J.J.D.), and Research Department of Computer Science (M.B.), University of Toronto, Canada; Department of Biostatistics (S.N.), Harvard Chan School of Public Health, Boston, MA; Department of Human Genetics (S.D.), University of Chicago, IL; and Techna Institute (M.B.), University Health Network, Toronto, Canada.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

The Article Processing Charge was funded by the authors.

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.

Glossary

RNA-seq = RNA sequencing; TCAG = The Centre for Applied Genomics; WES = whole exome sequencing; WGS = whole genome sequencing.

Methods

RNA-seq

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 RNeasy mini kit. Libraries for RNA-seq 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. RNA-seq data generation and analysis were carried out as described.⁶

Sequence Alignment, Expression, Variant Calling, and Splicing

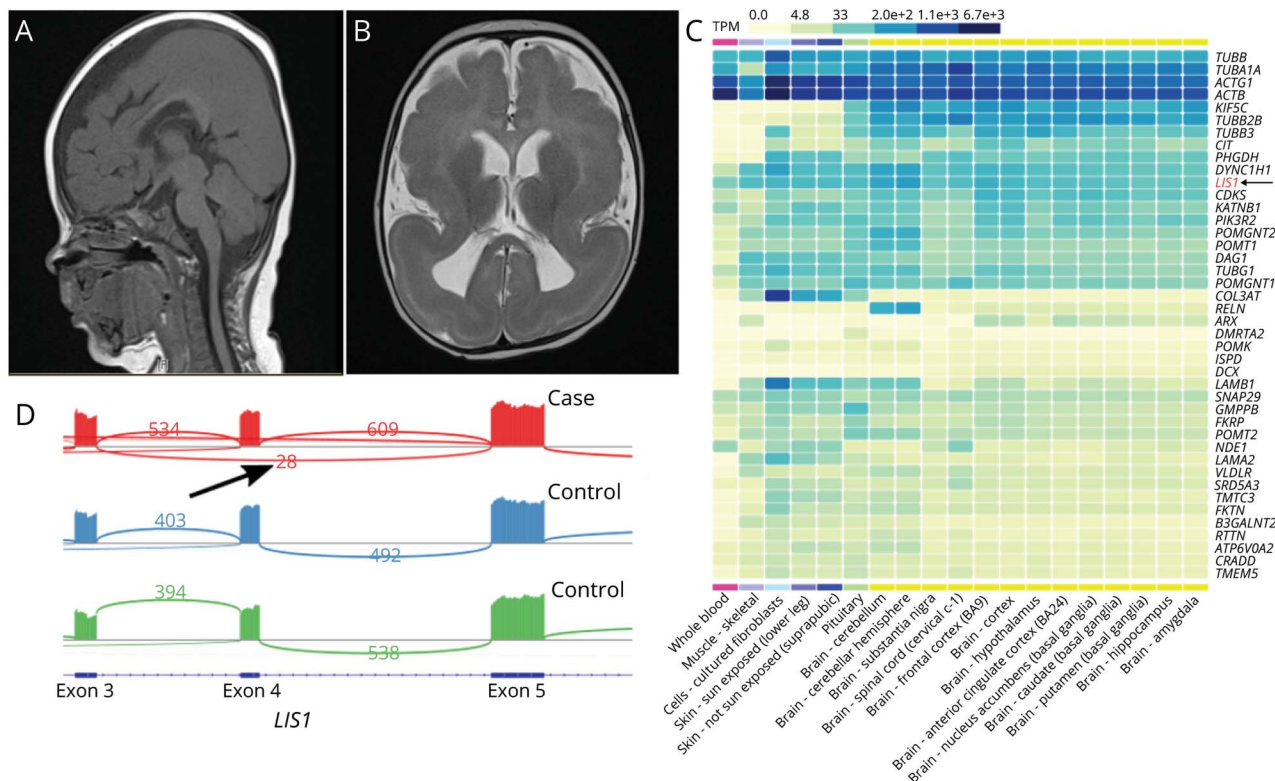
Alignment, variant calling, and quality controls steps were carried out using the RNA-seq workflow from the bc-bio-

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 multigene query portal (GTEx Portal) to generate a gene expression heatmap. For the 42 brain malformation genes (listed in the Figure, C, identified from the GeneDx panel), we selected tissues of interest (all brain tissues, fibroblasts, whole-blood, skeletal muscle, and skin) and

Figure Three-Year-Old Boy With Posterior-Anterior Gradient Lissencephaly



(A) Sagittal T1 MRI shows 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 shows agyria/pachygyria in the parietooccipital region. (C) Heat map coverage of common brain malformation genes in fibroblast vs blood vs brain tissue. Expression (in transcript 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 (6.7e + 3 TPM). Most of the 42 genes are expressed at higher levels in fibroblasts (column 3) compared to whole blood (column 1). In addition, the expression in fibroblast is more consistent with expression in brain tissues. (D) Sashimi plot shows transcript change with exon 4 skipping. We show expression of exons 3, 4, and 5 of *LIS1* in the proband (case) and 2 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.

exported svg (scalable vector graphics) of the resulting heatmap. The image was used to generate the Figure, C, where we highlight *LIS1* (the gene of interest).

Results

Based on the negative DNA testing, we pursued RNA sequencing through a research ethics board–approved research protocol. Because source material is a critical consideration for RNA-seq, and 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, C). 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 RNA-seq. The resulting total transcriptome analysis revealed reduced expression and allele imbalance of *LIS1*, with approximately 20 reads coming from 1 allele vs >500 from the other. Within the allele with reduced expression, we observed multiple splicing disruptions, including exon 4 skipping, although the low overall read count presented difficulty for the definitive interpretation of specific splicing alterations. Overall, the RNA-seq data were consistent with reduced expression of 1 *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 support the reduction of expression from this allele, and suggest 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 concluded that this heterozygous stop variant is the cause of disease in our patient.

Discussion

Our study demonstrates the suitability of RNA-seq for identifying novel mutations and for providing functional evidence of mutation consequence. Using RNA-seq, 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 RNA-seq 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 multigene panel or WES.

In terms of RNA-seq 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 noninvasive 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 levels (typically >1 transcript per million for expression and at least 5 reads for identifying splicing aberrations) can be meaningfully analyzed. Another potential source material for diseases of the CNS 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 or clarify pathogenic variants that cause Cornelia de Lange syndrome.⁸

In addition, 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 reasons why the variant was not detected are not clear, as overall mean read depth of the gene in the panel was >500×. 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 RNA-seq in the diagnostic pathway. Multigene panel or WES remain first-line testing methodologies for mutation identification across the spectrum of neurogenetic disorders. However, these technologies do not provide information related to the noncoding genome, and in addition often identify variants that are of uncertain significance. RNA-seq is therefore an ideal next step modality in panel or WES-negative cases or cases with variants that are of uncertain significance, as it provides data on variants (coding or noncoding) that affect 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 RNA-seq is also well positioned to provide functional annotation and interpretation when paired with WGS.

Acknowledgment

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

Study Funding

GenomeCanada DIG phase II (M.B., J.J.D.).

Disclosure

H. Qashqari, A. Ramani, and H. Gonorazky report no disclosures relevant to the manuscript. K. Amburgey is an employee at Deep Genomics. M.M. Ghahramani Seno reports no disclosures relevant to the manuscript. M. Brudno is an employee at Gene42 Inc. S. Naumenko, S. Das, and J.J. Dowling report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

Appendix Authors

Name	Location	Contribution
Hebah Qashqari, MBBS, FRCPC	Division of Neurology, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Arun Ramani, PhD	Centre for Computational Medicine, The Hospital for Sick Children, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Hernan Gonorazky, MD	Division of Neurology, Department of Pediatrics, and Program of Genetic and Genome Biology, The Hospital for Sick Children, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Kimberly Amburgey, MSc	Division of Neurology, Department of Pediatrics, and Program of Genetic and Genome Biology, The Hospital for Sick Children, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content

Appendix (continued)

Name	Location	Contribution
Mohammad M. Ghahramani Seno, DVM, PhD	Program of Genetic and Genome Biology, The Hospital for Sick Children, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Michael Brudno, PhD	Program of Genetic and Genome Biology and Centre for Computational Medicine, The Hospital for Sick Children, and Research Department of Computer Science, University of Toronto; and Techna Institute, University Health Network, Toronto, Canada	Analysis or interpretation of data
Sergey Naumenko, PhD	Department of Biostatistics, Harvard Chan School of Public Health, Boston, MA	Analysis or interpretation of data
Soma Das, PhD	Department of Human Genetics, University of Chicago, IL	Analysis or interpretation of data
James J. Dowling, MD, PhD	Division of Neurology, Department of Pediatrics, and Program of Genetic and Genome Biology, The Hospital for Sick Children, and Departments of Pediatrics and Molecular Genetics, University of Toronto, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

References

- Barkovich AJ, Dobyns WB, Guerrini R. Malformations of cortical development and epilepsy. *Cold Spring Harb Perspect Med*. 2015;5(5):a022392.
- Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA*. 2014;312(18):1880-1887.
- Rochtus A, Olson HE, Smith L, et al. Genetic diagnoses in epilepsy: the impact of dynamic exome analysis in a pediatric cohort. *Epilepsia*. 2020;61(2):249-258.
- Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet*. 2016;17(5):257-271.
- Cummings BB, Marshall JL, Tukiainen T, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med*. 2017;9(386).
- Gonorazky HD, Naumenko S, Ramani AK, et al. Expanding the boundaries of RNA sequencing as a diagnostic tool for rare mendelian disease. *Am J Hum Genet*. 2019;104(3):466-483.
- Dobyns WB, Das S. *PFAFH1B1-Associated lissencephaly/subcortical band heterotopia*. GeneReviews; 1993.
- Rentas S, Rathi KS, Kaur M, et al. Diagnosing Cornelia de Lange syndrome and related neurodevelopmental disorders using RNA sequencing. *Genet Med*. 2020;22(5):927-936.
- Lionel AC, Costain G, Monfared N, et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med*. 2018;20(4):435-443.