REVIEW

WILEY

RNA sequencing role and application in clinical diagnostic

Fatemeh Peymani[1,2] | Aiman Farzeen[1,2] | Holger Prokisch[1,2]

[1]School of Medicine, Institute of Human Genetics, Technical University of Munich, Munich, Germany
[2]Institute of Neurogenomics, Computational Health Center, Helmholtz Centre Munich, Neuherberg, Germany

Correspondence
Holger Prokisch, School of Medicine, Institute of Human Genetics, Technical University of Munich, Munich, Germany.
Email: prokisch@helmholtz-muenchen.de

Funding source
Bavarian State Ministry of Health and Care, Grant/Award Number: DMB-1805-0002

Received: 30 December 2021
Accepted: 18 January 2022

ABSTRACT

Although whole-exome sequencing and whole-genome sequencing has tremendously improved our understanding of the genetic etiology of human disorders, about half of the patients still do not receive a molecular diagnosis. The high fraction of variants with uncertain significance and the challenges of interpretation of noncoding variants have urged scientists to implement RNA sequencing (RNA-seq) in the diagnostic approach as a high throughput assay to complement genomic data with functional evidence. RNA-seq data can be used to identify aberrantly spliced genes, detect allele-specific expression, and identify gene expression outliers. Amongst eight studies utilizing RNA-seq, a mean diagnostic uplift of 15% has been reported. Here, we provide an overview of how RNA-seq has been implemented to aid in identifying the causal variants of Mendelian disorders.

KEYWORDS

Aberrant expression, Clinical diagnosis, Gene expression outliers, Genetics diagnosis, RNA phenotype, RNA sequencing, Transcriptome

INTRODUCTION

Accurate identification of causal variants in individuals with Mendelian disorders is of paramount importance for patient management. The identification of causal variants aids in providing genetic diagnosis, enabling prognostic guidance, family risk assessment, and offering a chance for personalized treatment.[1] In the past decade, next-generation sequencing (NGS) has revolutionized molecular diagnostics of rare Mendelian disorders. The diagnostic rates of whole-exome sequencing (WES), focus on coding sequences representing 1.5% of the human genome, ranging from 28% to 55%.[2] While whole-genome sequencing (WGS) provides comprehensive information on the 3 billion bases of the human DNA, the diagnostic yield over WES improves by only 5% (Figure 1).[3] This difference is minor when compared to the diagnostic yield from WES. The trivial difference reflects limitations in clinical interpretation of noncoding variants detected additionally by WGS. Coding variants make up more than 90% of the pathogenic/likely pathogenic variants in clinical databases.[4] Nevertheless, the noncoding variants have long been established to play an important role in human diseases.[5] The interpretation of those variants is challenging using genomic information alone.

The large number of variants identified by WES/WGS poses a new challenge of variant interpretation. To structure this process, the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) established a guideline, incorporating 28 criteria for variant classification. According to these

DOI: 10.1002/ped4.12314

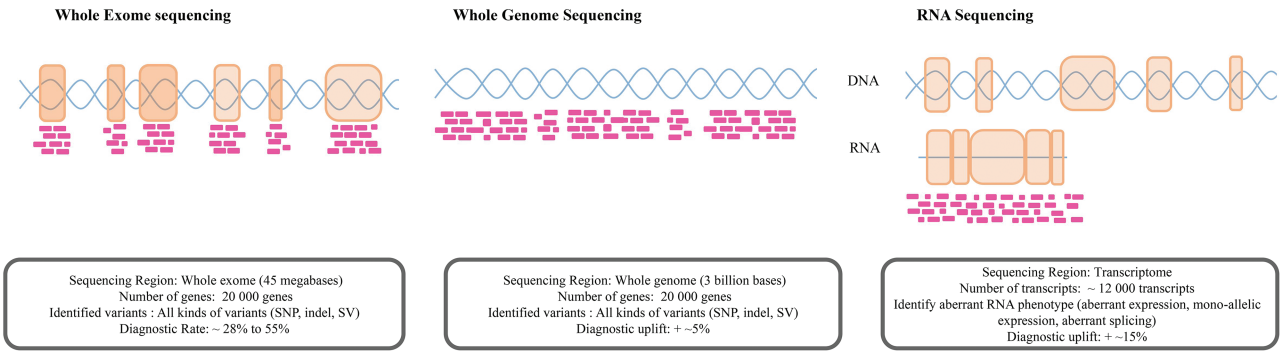

FIGURE 1 Coverage, identified variants and diagnostic rate of whole-exome sequencing, whole-genome sequencing, and RNA sequencing.

criteria, which take advantage of population data, functional data, segregation data, computational data, *de novo* data, and allelic data, the variants are classified into five categories: "pathogenic", "likely pathogenic", "uncertain significance", "likely benign", and "benign".[6]

The largest fraction of variants is classified as variants of uncertain significance (VUS). The limited knowledge of the functional consequences of rare genomic variants, majorly noncoding variants, hampers the diagnostic rate of WES and WGS to exceed 50%. Since 2017, RNA sequencing (RNA-seq) has been used as a complementary tool to DNA-based sequencing for genetic diagnosis of Mendelian disorders and it has been shown to improve the diagnostic yield over WES by about 15% (Figure 1).[7,8] RNA-seq enables the identification of aberrant RNA phenotypes which include aberrant gene expression, mono-allelic expression, and aberrant splicing. Therefore, RNA-seq improves diagnosis through both illuminating the functional consequence of VUS and helping in the prioritization of variants neither prioritized by WGS nor detected by WES. In this review, we provide an overview of aberrant RNA phenotypes and the studies utilizing RNA-seq as a complementary tool to DNA sequencing (DNA-seq) techniques.

DEFINITION OF THE TRANSCRIPTOME

The transcriptome can be defined as a complete set of intracellular transcripts and the amount of those transcripts for a particular developmental stage or physiological state.[9] Compared to DNA, which remains nearly identical across all cells and time, the transcriptome is very dynamic. The transcriptome acts as a transient intermediary between DNA and proteins.[10] The whole transcriptome of a cell can be sequenced and analyzed in a single run by RNA-seq.[11] RNA-seq data provide comprehensive information on the sequence, structure, and quantity of specific RNA sequences. Therefore, RNA-seq is regarded as an invaluable tool for the systematic detection of RNA phenotypes. Pathological RNA phenotypes can be the result of rare DNA variants.[7,8] Hence, RNA-seq is applied to screen for variants that cause extreme RNA phenotypes, that is, aberrant expression, aberrant splicing, and mono-allelic expression (Figure 2).[7] These three aberrant RNA phenotypes and how we can take advantage of them in identifying putative disease-causing variants are explained in detail in the following paragraphs.

Aberrant expression

Aberrant expression is defined as an expression that significantly deviates from the normal physiological range. A number of processes can affect gene expression. Chromatin packing, histone modification, transcription initiation, RNA polyadenylation, splicing, and translation initiation all tightly regulate gene expression. Genetic variation, both coding and noncoding, is well known to impact these processes.[12] Low expression level of a disease gene often translates into low protein level and indicates a genetic disorder.

The promoter region of a gene is a crucial component in the initiation and regulation of gene expression. However, promoter regions are not so well conserved. It is estimated that 1% of disease-causing variants are located within promoter regions. Those variants cause disease through their effects on gene transcription. Low transcript levels can also be explained by RNA degradation. Nonsense-mediated mRNA decay (NMD) is likely to result in a very low expression level. Nonsense, frameshift, and often canonical splice site variants, can trigger NMD by introducing premature termination codons. The severity of the reduced expression levels correlates with the number of affected alleles and the effect size of splice defects.[7]

Gene expression is measured by quantifying read counts that are mapped to a gene. Read counts should be normalized for sample sequencing depth and gene length. Gene expression outliers are detected as reading counts that significantly deviate from the read count distribution of the reference dataset. Different methods have been applied for

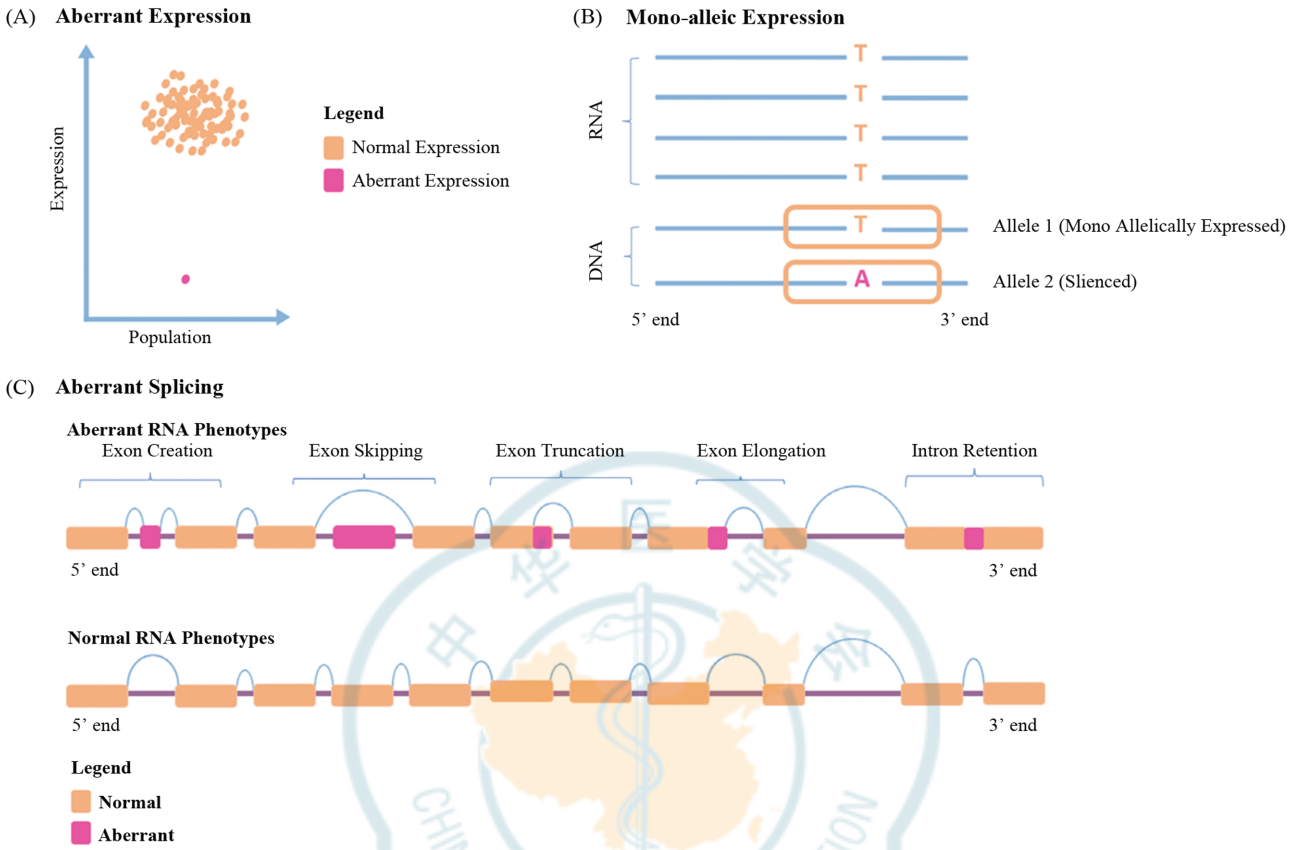

FIGURE 2 Three aberrant RNA phenotypes identified by RNA sequencing (RNA-seq). (A) Aberrant expression, (B) Mono-allelic expression, and (C) Aberrant splicing versus normal slicing.

outlier detection including DESeq2,[13] *Z*-score approach,[7] and OUTRIDER (OUTlier in RNA-seq fInDER).[14] The hallmarks of outrider include the correcting for gene covariation resulting from technical, environmental, or common genetic variations.[14] When using fibroblast cell lines, on average, five outliers are called in each individual. Combined visual inspection of RNA-seq and genomic data of outlier genes is necessary for the identification of pathogenic variants.[15]

Yépez et al.[15] demonstrated the application of RNA-seq in diagnostics. In this study, RNA-seq analysis in a patient, whose WES analysis was inconclusive, revealed two underexpression outliers. Amongst those outliers, one gene called *UFM1* was phenotypically matched patient phenotype and showed the lowest expression in the sample compared to the other 303 samples. WES reinspection identified a 3-bp homozygous deletion in the promoter region of the *UFM1* gene. Of note, the variant had been independently reported to negatively affect promoter and transcription activity (Figure 3). Overall, Yépez et al.[15] provided a genetic diagnosis in about 10% of WES/WGS unsolved cases through detection of aberrant gene expression.

Mono-allelic expression

RNA-seq is also used for the detection of mono-allelic expression. Mono-allelic expression (MAE), is referred to the situation when mainly one of the two alleles is expressed, while the second allele is only expressed at very low levels (Figure 2B). This phenomenon may be the result of an allele being epigenetically silenced, not expressed due to promoter variants, or an allele being degraded. Some common examples of epigenetic silence of alleles include X-inactivation and genomic imprinting.[16] Heterozygous null variants triggering NMD are examples of allele-specific degradation. As a result, variants located on trans are mono-allelically expressed. Promoter variants are usually missed by WES and skewed X-inactivation resulting in MAE can also not be detected by genome sequencing.

Often a single heterozygous rare variant escapes prioritization in disease with a recessive mode of inheritance. When only one allele carrying a rare deleterious variant is expressed, and the other allele is nearly completely silenced, the variant behaves like a homozygous variant on the transcript level. This phenomenon can be detected

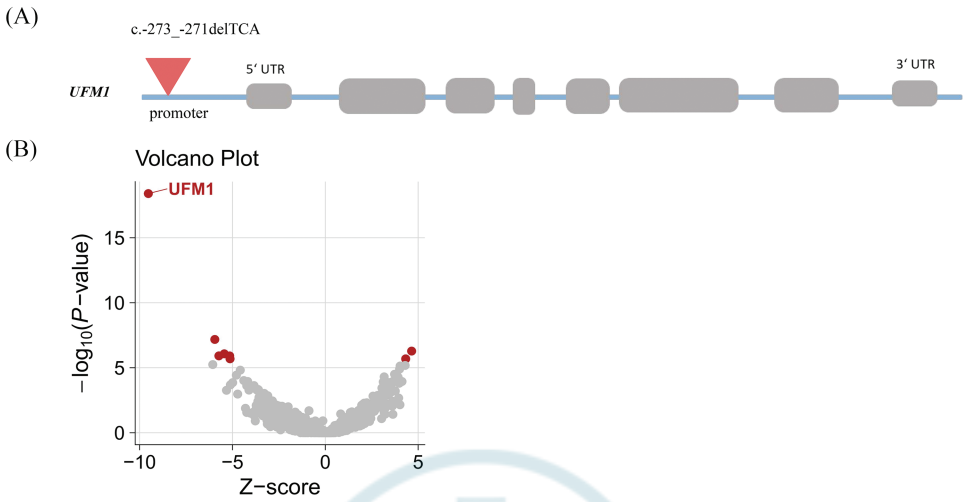

FIGURE 3 Identification of the causative variant with the help of aberrant expression detection. (A) 3 bp deletion in the promoter of *UFM1* gene (red triangle) (NM_016617.2: c. −273_−271delTCA) is reprioritized in whole-exome sequencing (WES) data after detection of *UFM1* aberrant expression by RNA sequencing (RNA-seq). (B) Gene-level distribution of $-\log_{10}$ (*P*-value) versus *Z*-score in the sample of interest (Yépez 2021).[15] Red dots represent outliers. *UFM1* is among the outliers with the lowest *Z*-score and *P*-value. WES, whole-exome sequencing; UTR, untranslated region.

in RNA profiles.[7] Currently, two methods have been developed for MAE detection, negative binomial tests, and ANEVA-DOT.[17] Given ANEVA-DOT is not applicable for all online Mendelian inheritance in man (OMIM) genes, negative binomial tests are more frequently used for MAE detection. Heterozygous single nucleotide variants need to be covered at least 10 times by RNA-seq for calling MAE. Allelic expression is detected by counting the reads aligned to each allele at genomic positions of heterozygous variants. Alleles with a significant allele expression ratio greater than 0.8 or lower than 0.2 can be considered to be mono-allelically expressed.[7] In about 2% of WES/WGS unsolved cases, MAE of a disease gene can be detected leading to a molecular diagnosis.

Aberrant splicing

It has been estimated that ~94% of human genes undergo splicing and in the majority of them, alternative splicing (AS) occurs. AS is regarded as a key cellular process in ensuring functional complexity in higher eukaryotes.[18] Defects in the splicing events have been recognized as the main cause of Mendelian disorder.[19–21] Splicing could be affected by both exonic and intronic variants. It is estimated that at least 10% of pathogenic variants have an effect on RNA splicing. The variants in canonical splice sites, located in ±1 or 2 exon-intron boundaries, are given a very strong level of pathogenicity according to the ACMG guideline.[6] Nevertheless, the effect of deep intronic, exonic, and synonymous variants on splicing is hard to be predicted based on the DNA sequence. Indeed, 25% of the rare synonymous variants can cause aberration in splicing.[20,21] The effect of synonymous and missense variants on splicing is often overlooked in clinical interpretation of WES data. At the same time, deep intronic variants are rarely prioritized as pathogenic in WGS. While in silico prediction tools, such as GeneSplicer,[22] SPANR,[23] and VEP,[24] work well in predicting the effect of variants in canonical splice sites, they often fail to accurately predict the effect of synonymous, missense, and deep intronic variants on splicing.[25] Therefore, RNA-seq plays an important role in directly probing splice isoforms and adding a functional layer to the variants detected by WES and WGS. Splice defects could result in exon skipping, exon truncation, exon elongation, exon creation, and intron inclusion (Figure 2C). Aberrant splicing is detected by comparing normalized split reads, spanning splice sites of two exons, between a case and reference dataset. Various methods such as FRASER,[26] LeafCutter/LeafCutterMD,[27] and SPOT[28] have been developed to systematically detect the aberrant splicing in RNA-seq data.

As an example, Kernohan et al.[29] performed whole blood RNA-seq on a patient affected by spinal muscular atrophy for whom DNA-seq reported only one pathogenic variant. Comparing the transcriptome profile of the affected individual with 909 controls detected skipping of exon 6. Exon skipping occurred as a result of a predicted missense variant which was located 2 bp from the splice junction.[29] Bioinformatic tools failed to predict the pathogenic effect of this missense variant on splicing function (Figure 4).

DATA ANALYSIS

Yepez et al.[30] provided a detailed semi-automated workflow for aberrant RNA phenotype identification called DROP (detection of RNA outliers pipeline).[30] DROP incorporates recent tools, providing an all-in-one

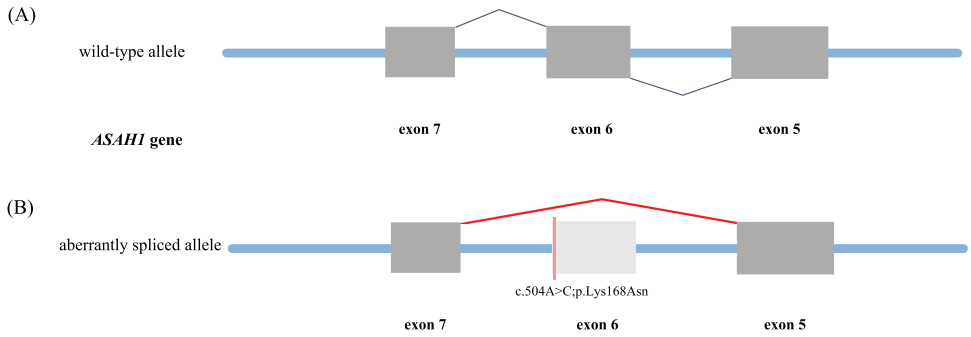

FIGURE 4 Exon skipping as a result of missense variant 2 bases from the splice junction. (A) Splicing pattern in wild-type allele. (B) The c.504A>C variant in the *ASAH1* gene leads to aberrant splicing (exon 6 skipping) and is regarded as a disease-causing variant.

computational workflow, for identification of aberrant expression, aberrant splicing, and mono-allelic expression. DROP standardizes all the required steps for achieving results from BAM files of RNA-seq and VCF files of DNA-seq as input and produces gene expression outliers, visualization of results, and browsable web page reports.

STUDIES UTILIZING RNA-SEQ

RNA-seq was utilized in diagnostics for the first time in 2017.[7,8] Since then, several studies performed RNA-seq on different tissues including fibroblast, muscle, and blood. The tissues were taken from individuals with diverse disorders, such as mitochondrial, muscular, and neurodevelopmental disorders. The studies implemented different methods and protocols and highlighted the diagnostic utility of RNA-seq. On average, among 544 undiagnosed cases, 94 received a diagnosis with the help of RNA-seq reflecting a diagnostic yield of 17%. In the following paragraph, an overview of the studies utilizing RNA-seq as a diagnostic tool is provided.

Kremer et al.[7] utilized RNA-seq to detect the genetic cause of patients with mitochondrial disorders who remained undiagnosed after WES and WGS. Transcriptome analysis of patients' fibroblasts led to improvement in diagnostic yield by 10% (5/48). Cummings et al.[8] applied RNA-seq on skeletal muscle samples, some with potential splice site variants, who suffered from suspected genetic muscular disorders. For these patients, WES and WGS were inconclusive. With incorporating RNA-seq, a diagnostic yield of 35% (17/50) was achieved. Gonorazky et al.[1] analyzed RNA-seq data from muscle biopsies, fibroblasts, and T-myotubes in individuals with neuromuscular disorders. This study resulted in the detection of aberrant RNA phenotype in nine out of 25 cases.[1] Frésard et al.[31] utilized blood transcriptome and reported causal variants in 17% of individuals affected by 16 different disease categories. Lee et al.[32] applied WES, WGS, and RNA-seq on 48 individuals with a diverse range of rare diseases and reported the genetic etiology in seven cases. By focusing on aberrant splicing in multiple tissues (whole blood, skin fibroblasts, and muscle), a diagnostic yield of 15% was achieved. Rentas et al.[33] performed RNA-seq on human B-lymphoblastoid cell lines from five patients with a neurodevelopmental disorder. In three out of five cases, a diagnosis was achieved by focusing on aberrant splicing. Murdock et al.[34] reported a diagnostic yield of 17% (14/83) when performing RNA-seq in blood or fibroblasts of WES negative cases suffering from various rare disorders. Most recently, Yepez et al.[15] applied RNA-seq on fibroblasts from a large cohort of WES or WGS undiagnosed patients. The RNA-seq analysis unraveled the molecular etiology in 16% (33/205) of the cases. Altogether, a number of different tissues have been successfully used for RNA-seq. Most studies focused on splicing. The diagnostic yield was generally around 15%. In light of all the studies, RNA-seq has been shown as an invaluable tool to improve molecular diagnostics.

TISSUE SPECIFICITY AND SAMPLE SIZE

While DNA-seq analysis is mainly tissue independent, RNA-seq analysis has to take tissue-specific expression into consideration. Different tissues or cell types represent a wide spectrum of splicing events and expression patterns. Selecting an appropriate tissue for RNA-seq plays an important role in obtaining the optimal diagnostic yield. Finding the right tissue with specific gene expression vs. accessibility and invasiveness of tissue collection is a known obstacle. Although blood is regarded as an easily accessible tissue, covering ~50% of OMIM genes, it's not always the best tissue to use. Murdock et al.[34] indicated that some pathogenic splicing variants detected in fibroblast remained undetected in blood. Yepez et al.[15] highlighted the effectiveness of muscle biopsy and fibroblast in RNA-seq studies, as they show better coverage (~70%) for OMIM genes. The affected tissue likely ensures that the gene of interest is expressed, however, secondary pathologies may mask the primary defect. This is less likely for splice defects but applies to aberrant expression.

The analysis of RNA-seq for diagnostic purposes requires normalized RNA-seq samples. Minimum of 30 and 50 cases for aberrant splicing and aberrant expression detection are recommended, respectively.[14,26] In addition, a public RNA-seq dataset will improve the statistical analysis.

The Genotype-Tissue Expression database is a comprehensive reference database, consisting of omics data (WGS, WES, and RNA-seq) of 54 nondiseased tissues across approximately 1000 individuals.[35] This database provides an invaluable resource for tissue selection and could also be used as a comprehensive control dataset for statistical comparison. However, the batch effect and technical bias of the NGS should be taken into consideration while using public data. The various mathematical methods such as denoising autoencoder[14] and principal component analysis could be applied to control for any possible biases.

CONCLUSION

While WES and WGS drastically increase our understanding of the genetic etiology of Mendelian disorders, around half of the patients remain undiagnosed. RNA-seq using blood, fibroblasts, and muscle biopsies has proven itself as a promising tool in clinical practice. It has led to about 15% diagnostic uplift on average based on the studies done so far. RNA-seq demonstrates success in prioritizing and detecting not only deleterious deep intronic variants but also coding variants affecting gene expression, which are often overlooked by WES and WGS. Information obtained on functional consequences of these variants could further be used to improve the performance of in silico prediction tools.

Even after employing RNA-seq, a larger fraction of patients remain undiagnosed. This indicates the necessity of integrating additional omics data such as proteomics, especially in cases where the causal variant does not lead to transcriptome aberration.[33] Furthermore, the reported studies utilized arbitrary thresholds for assessing aberrant RNA phenotypes, and no consensus exists in this regard. Therefore, improved analytical tools and appropriate guidelines for interpreting aberrant RNA phenotypes are needed to be incorporated into the ACMG/AMP guidelines.[21]

CONFLICT OF INTEREST

There is no conflict of interest.

REFERENCES

1. Gonorazky HD, Naumenko S, Ramani AK, Nelakuditi V, Mashouri P, Wang P, et al. Expanding the boundaries of RNA sequencing as a diagnostic tool for rare Mendelian disease. *Am J Hum Genet*. 2019;104:466-483. DOI: 10.1016/j.ajhg.2019.01.012
2. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. 2016;18:696-704. DOI: 10.1038/gim.2015.148
3. Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, et al. 100,000 genomes pilot on rare-disease diagnosis in health care–preliminary report. *N Engl J Med*. 2021;385:1868-1880. DOI: 10.1056/NEJMoa2035790
4. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42:D980-D985. DOI: 10.1093/nar/gkt1113
5. Zhang F, Lupski JR. Non-coding genetic variants in human disease. *Hum Mol Genet*. 2015;24:R102-R110. DOI: 10.1093/hmg/ddv259
6. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405-424. DOI: 10.1038/gim.2015.30
7. Kremer LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun*. 2017;8:15824. DOI: 10.1038/ncomms15824
8. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med*. 2017;9:eaal5209. DOI: 10.1126/scitranslmed.aal5209
9. Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. *PLoS Comput Biol*. 2017;13:e1005457. DOI: 10.1371/journal.pcbi.1005457
10. Cieślik M, Chinnaiyan AM. Cancer transcriptome profiling at the juncture of clinical translation. *Nat Rev Genet*. 2018;19:93-109. DOI: 10.1038/nrg.2017.96
11. Wang Z, Gerstein M, Snyder M. RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57-63. DOI: 10.1038/nrg2484
12. Curry PDK, Broda KL, Carroll CJ. The role of RNA-sequencing as a new genetic diagnosis tool. *Curr Genet Med Rep*. 2021;9:13-21. DOI: 10.1007/s40142-021-00199-x
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. DOI: 10.1186/s13059-014-0550-8
14. Brechtmann F, Mertes C, Matusevičiūtė A, Yépez VA, Avsec Z, Herzog M, et al. OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am J Hum Genet*. 2018;103:907-917. DOI: 10.1016/j.ajhg.2018.10.025
15. Yépez VA, Gusic M, Kopajtich R, Mertes C, Smith NH, Alston CL, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *medRxiv*. Published online April 5, 2021. DOI: 10.1101/2021.04.01.21254633v1

16. Bjornsson HT. The Mendelian disorders of the epigenetic machinery. *Genome Res*. 2015;25:1473-1481. DOI: 10.1101/gr.190629.115
17. Mohammadi P, Castel SE, Cummings BB, Einson J, Sousa C, Hoffman P, et al. Genetic regulatory variation in populations informs transcriptome analysis in rare disease. *Science*. 2019;366:351-356. DOI: 10.1126/science.aay0256
18. Chen L, Tovar-Corona JM, Urrutia AO. Alternative splicing: a potential source of functional innovation in the eukaryotic genome. *Int J Evol Biol*. 2012;2012:596274. DOI: 10.1155/2012/596274
19. Tazi J, Bakkour N, Stamm S. Alternative splicing and disease. *Biochim Biophys Acta*. 2009;1792:14-26. DOI: 10.1016/j.bbadis.2008.09.017
20. Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med*. 2012;18:472-482. DOI: 10.1016/j.molmed.2012.06.006
21. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet*. 2016;17:19-32. DOI: 10.1038/nrg.2015.3
22. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res*. 2001;29:1185-1190. DOI: 10.1093/nar/29.5.1185
23. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RKC, et al. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science*. 2015;347:1254806. DOI: 10.1126/science.1254806
24. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17:122. DOI: 10.1186/s13059-016-0974-4
25. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res*. 2014;42:13534-13544. DOI: 10.1093/nar/gku1206
26. Mertes C, Scheller IF, Yépez VA, Çelik MH, Liang Y, Kremer LS, et al. Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat Commun*. 2021;12:529. DOI: 10.1038/s41467-020-20573-7
27. Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, et al. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet*. 2018;50:151-158. DOI: 10.1038/s41588-017-0004-9
28. Ferraro NM, Strober BJ, Einson J, Abell NS, Aguet F, Barbeira AN, et al. Transcriptomic signatures across human tissues identify functional rare genetic variation. *Science*. 2020;369:eaaz5900. DOI: 10.1126/science.aaz5900
29. Kernohan KD, Frésard L, Zappala Z, Hartley T, Smith KS, Wagner J, et al. Whole-transcriptome sequencing in blood provides a diagnosis of spinal muscular atrophy with progressive myoclonic epilepsy. *Hum Mutat*. 2017;38:611-614. DOI: 10.1002/humu.23211
30. Yépez VA, Mertes C, Müller MF, Klaproth-Andrade D, Wachutka L, Frésard L, et al. Detection of aberrant gene expression events in RNA sequencing data. *Nat Protoc*. 2021;16:1276-1296. DOI: 10.1038/s41596-020-00462-5
31. Frésard L, Smail C, Ferraro NM, Teran NA, Li X, Smith KS, et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med*. 2019;25:911-919. DOI: 10.1038/s41591-019-0457-8
32. Lee H, Huang AY, Wang LK, Yoon AJ, Renteria G, Eskin A, et al. Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet Med*. 2020;22:490-499. DOI: 10.1038/s41436-019-0672-1
33. Rentas S, Rathi KS, Kaur M, Raman P, Krantz ID, Sarmady M, et al. Diagnosing Cornelia de Lange syndrome and related neurodevelopmental disorders using RNA sequencing. *Genet Med*. 2020;22:927-936. DOI: 10.1038/s41436-019-0741-5
34. Murdock DR, Dai H, Burrage LC, Rosenfeld JA, Ketkar S, Müller MF, et al. Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of conventional genomic testing. *J Clin Invest*. 2021;131:e141500. DOI: 10.1172/JCI141500
35. GTEx Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet*. 2013;45:580-585. DOI: 10.1038/ng.2653

How to cite this article: Peymani F, Farzeen A, Prokisch H. RNA sequencing role and application in clinical diagnostic. *Pediatr Investig*. 2022;6:29–35. https://doi.org/10.1002/ped4.12314