Journal of Translational Medicine

RESEARCH

Open Access

Blood RNA-seq in rare disease diagnostics: a comparative study of cases with and without candidate variants

Xiaomei Luo[1], Bing Xiao[1], Lili Liang[2], Kaichuang Zhang[2], Ting Xu[1], Huili Liu[1], Yi Liu[1], Yongguo Yu[1,2*] and Yanjie Fan[1*]

Abstract

Background Approximately 60% of rare disease cases remain unsolved after exome and genome sequencing (ES/GS). Blood RNA sequencing (RNA-seq) complements DNA-level diagnosis by revealing the functional impact of variants on gene expression and splicing, but to what extent RNA-driven approaches offer diagnostic benefits across different scenarios—with and without pre-existing candidate variants—remains uncertain.

Methods 128 unrelated probands with suspected Mendelian disorders who had previously undergone ES/GS were recruited. A validation cohort (n = 7, with variants expected to alter RNA) and a test cohort (n = 121, including 10 with variants of uncertain significance (VUS) and 111 with no previously identified candidate variants) were analyzed. Blood RNA-seq was performed, and aberrant splicing (AS) and aberrant expression (AE) were detected using the DROP pipeline. SpliceAI predictions were compared with RNA-seq results for splicing-related VUS variants, and pathogenicity was re-evaluated. AS/AE outliers were evaluated for diagnostic potential in cases without candidate variants. The feasibility of an RNA-driven approach was assessed by ranking causal variant-associated aberrant events.

Results The pipeline correctly identified all expected AS/AE events in the validation cohort. In the test cohort with candidate VUS, RNA-seq provided a 60% (6/10) diagnostic uplift. Notably, SpliceAI predictions matched RNA-seq observations perfectly only in 40% of these VUS. A 2.7% (3/111) diagnostic uplift was achieved in the test cohort with no prior candidates. Overall, target AS and AE events ranked among the top eight in 14 of the 16 diagnosed cases using a purely RNA-driven approach; however, two cases would have been missed without prior candidate identification from DNA sequencing.

*Correspondence:
Yongguo Yu
yuyongguo@shsmu.edu.cn
Yanjie Fan
fanyanjie@shsmu.edu.cn
[1] Clinical Genetics Center, Shanghai Institute for Pediatrics, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Room 801, Science and Education Building, No.1665, Kong Jiang Road, Shanghai 200092, China
[2] Department of Pediatric Endocrinology and Genetic Metabolism, Shanghai Institute for Pediatrics, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 20092, China

Conclusion Blood RNA-seq is highly effective in refining the interpretation of splicing VUS, frequently leading to reclassification and diagnosis. Meanwhile, RNA-driven identification of causal variants shows a more modest yield in cases without prior candidates. This study supports an RNA-complementary approach as the preferred strategy for clinical utility.

Keywords RNA sequencing, Undiagnosed rare disease, Variant interpretation, Genome sequencing, Clinical practice

Introduction

Despite the increasing application of exome sequencing (ES) and genome sequencing (GS) to identify the molecular causes of rare diseases, a significant challenge remains: approximately 60% of cases remain unsolved [1, 2]. With the increasing ability to detect variants of complex types and in non-coding regions, the diagnostic gap exists in fully elucidating the functional significance of detected variants. Currently unsolved cases after ES/GS can largely be attributed to: (i) variants of uncertain significance (VUS) in genes potentially related to patient phenotypes; (ii) lack of clear candidate variants. These challenges hinder clinical decision making and require complementary approaches.

To overcome the limitations of DNA-centric variant detection and interpretation, RNA sequencing (RNA-seq) has emerged as a powerful complementary tool. By directly assessing the transcriptome, RNA-seq can reveal the functional consequences of both coding and non-coding genetic variants on gene expression and splicing [3]. Aberrant RNA processing, including altered expression levels and splicing defects, is increasingly recognized as a key mechanism in genetic disorders [4]. Pioneering studies indicate that RNA-seq increases diagnostic rates by 8–36% over DNA sequencing alone, depending on the disease and tissue type [3, 5–9]. While tissue accessibility can be a limiting factor for comprehensive RNA analysis, blood RNA-seq offers a feasible and minimally invasive approach. Blood RNA analysis has been utilized for both VUS assessment in the presence of candidate variants [10–12] and de novo discovery in the absence of candidate variants [13, 14]. However, the extent to which RNA-driven approaches offer diagnostic advantages in different scenarios—with and without pre-existing candidate variants—remains to be determined in order to guide appropriate clinical use.

In this study, we investigate the clinical value of blood RNA-seq in a cohort of rare disease patients with heterogeneous phenotypes who remain undiagnosed after ES or GS. Specifically, we aim to determine: (i) the additive diagnostic utility of blood RNA-seq in cases with pre-existing candidate variants, particularly splicing-related VUS, and to which extent the RNA data refine variant interpretation and diagnostic outcomes; (ii) the diagnostic uplift and feasibility of a blood RNA-driven approach in cases lacking prior candidate variants or genes, evaluating the recall of relevant variants and the practical workload for candidate identification. Our findings will delineate the preferred strategy to use blood RNA-seq – whether as an RNA-driven first-line approach or as an RNA-complementary strategy following ES/GS—thus inform best practices for incorporating RNA-seq into the diagnostic flow for rare diseases.

Materials and methods

Study design and cohort recruitment

All individuals were referred and evaluated by clinicians from various specialties at Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. This study was conducted in accordance with the principles of the Declaration of Helsinki. Approval was obtained from the Ethics Committee of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (XHCC-C-2022–077-1). Written informed consent for RNA-seq testing and publication was obtained from the patients or their legal guardians.

The study design is shown in Fig. 1A. A total of 128 unrelated probands with suspected Mendelian disorders were recruited for this study, all of whom had previously undergone ES or GS testing, including two cohorts: a validation cohort (n = 7, ES/GS solved, used for method and analysis validation) and a test cohort (n = 121, ES/GS unsolved cases). Detailed phenotypic information and previous ES or GS results are provided in Supplementary Table 1.

The validation cohort consisted of cases with variants deemed to alter splicing (n = 3, single nucleotide variants in canonical splice sites or with high confidence by SpliceAI prediction) or variants deemed to alter expression (n = 4, consisted of three copy number variations (CNVs, including an autosomal deletion, a deletion in chromosome X, a triplication) and one small indel variant triggering nonsense-mediated mRNA decay (NMD)). Variant information, aberrations detected by RNA-seq and altered production description of the validation cohort are listed in Supplementary Table 2.

The test cohort includes 10 cases with candidate splicing VUS ("splicing VUS" cohort) and 111 cases without

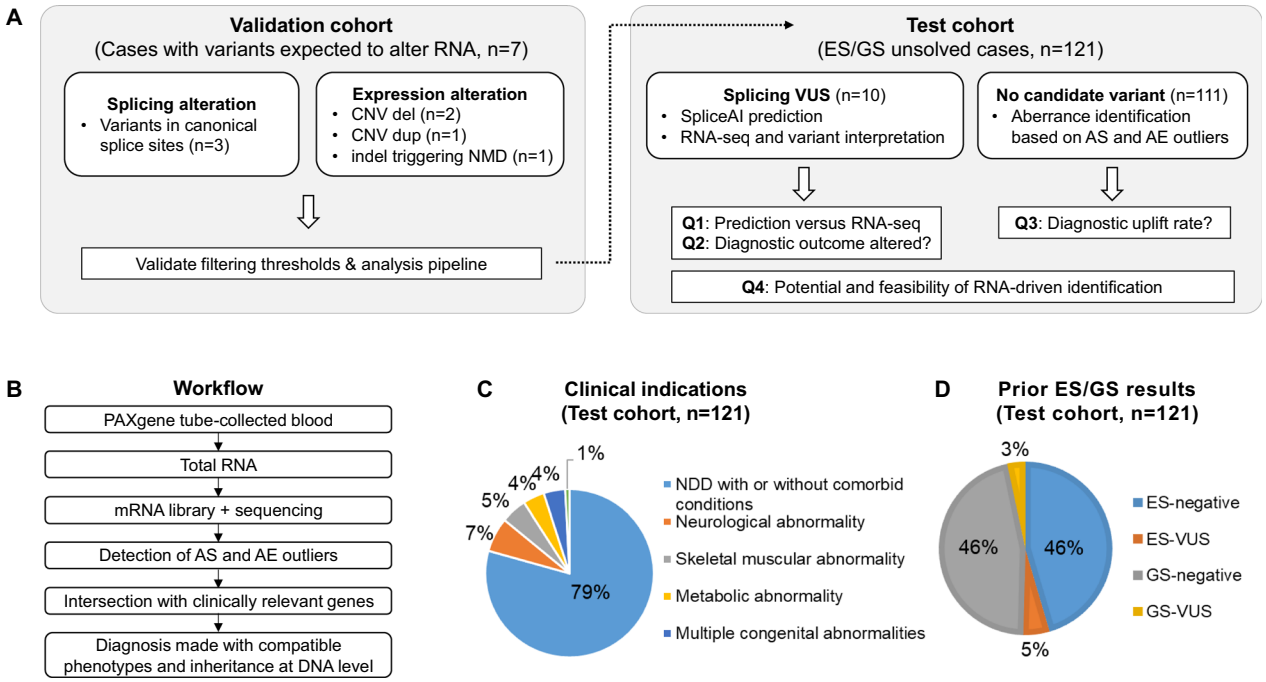

Fig. 1 Overview of the study design, basic information about the test cohort, and RNA-seq workflow. **A** Study design and questions to be addressed. **B** RNA-seq workflow in this study. **C** Clinical indications of the test cohort. **D** Prior ES or GS results of the test cohort

previously identified candidate variants ("No candidate variant" cohort) (Fig. 1A). Cases in the "splicing VUS" cohort were recruited from unsolved ES/GS cases meeting the following criteria: (i) VUS were identified in genes associated with diseases consistent with the patient's phenotype and inheritance pattern; (ii) variants potentially affect splicing with a SpliceAI prediction score above 0.2 [15]; and (iii) the interrogated VUS was in a gene expressed in blood above a minimum level (TPM > 1 in this study). The predicted mRNA consequence by SpliceAI was compared with RNA-seq results in this study. The pathogenicity of candidate VUS was reclassified according to the ClinGen SVI Splicing Subgroup recommendations combining evidence from RNA-seq data [16], and the diagnostic outcome was re-evaluated. The "No candidate variant" cohort includes cases where ES/GS did not identify any variants considered explanatory for the patient's phenotype, not even VUS. For this cohort, aberrant splicing (AS) and expression (AE) events in clinically relevant genes were identified in an RNA-driven approach and then cross-referenced with DNA variants in corresponding genes (Fig. 1B). The diagnostic uplift rate was assessed, and the potential and feasibility of RNA-driven candidate identification in the test cohort was further evaluated.

The clinical indications of patients in the test cohort (n = 121) are shown in Fig. 1C, with the majority (79%) of patients manifesting neurodevelopmental disorders with or without comorbid conditions, and the remaining 21% of patients manifesting neurological, skeletal muscular, metabolic, multiple congenital anomalies or other symptoms. Prior to RNA-seq, 60 samples underwent GS (4 in "splicing VUS" cohort, 56 in "No candidate variant" cohort); and 61 samples underwent ES only (6 in "splicing VUS" cohort and 55 in "No candidate variant" cohort). The composition of previous ES/GS results in the test cohort is shown in Fig. 1D.

RNA sequencing and data analysis

The RNA-seq workflow from sample collection to data analysis is shown in Fig. 1B. Whole blood was collected in a PAXgene Blood RNA tube (BD Biosciences), and total RNA was extracted using the PAXgene Blood RNA kit (Qiagen, Germany). Oligo dT-enriched mRNA was used for subsequent library preparation using the NEBNext Globin and rRNA Depletion Kit and NEBNext Ultra Directional RNA Library Prep Kit (NEB, Massachusetts, USA), followed by sequencing on the Novaseq 6000 instrument (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end reads. About 100M reads were generated per sample. Reads were aligned to the GRCh37 release 75 reference (to maintain consistency with the clinical laboratory exome/genome pipeline) using STAR V.2.7.9.a in two-pass mode [17]. Alignment files were utilized for outlier detection or visualized in the Integrative Genomics Viewer software (https://igv.org/) [18]. Quality control

was performed with RSeQC [19]. Outlier detection was performed using DROP V.1.4.0 with the recommended settings for AE and AS detection [20], see Supplementary Table 3 for specific parameters. Detected events were processed using an in-house pipeline to annotate disease association and functional consequences.

Analysis of cases in splicing VUS cohort

DNA-based splicing prediction

SpliceAI prediction was used (https://spliceailookup.broadinstitute.org/), and a threshold of $\Delta score \geq 0.2$ was used to predict AS [16]—described as "aberrance predicted".

RNA-based splicing analysis

Aberrant splicing (AS) involving junctions related to the VUS was considered "AS detected" if either of the following criteria was met: (i) Based on the FRASER2 output, the event showed $|\Delta\psi| \geq 0.2$ and a nominal p-value < 0.05. The nominal p-value was used here, rather than an adjusted p-value, because the strong prior evidence from the implicated VUS; (ii) Based on IGV inspection of splicing patterns around the VUS, at least 15 reads supported mis-splicing (threshold adapted from Australasian Consortium for RNA Diagnostics) [10]. "Consequence match" was defined as not only the detection of AS events, but also the concordance in the mRNA production—such as exon skipping or intron retention, or a mixture of both—between SpliceAI predictions and empirical observation in RNA-seq results.

RNA-based expression analysis

For splicing VUS, aberrant expression (AE) detection was only considered when no AS was detected – since abnormal junctions or transcripts may be efficiently removed by NMD, thereby masking splicing outliers. For VUS without detectable AS, the OUTRIDER module of DROP pipeline was used for expression analysis. Altered expression was defined by a p-value < 0.05, and fold change was reported to indicate the magnitude of dysregulation.

NMD prediction: "NMD triggering or escape" is based on the premature termination codon (PTC) not occurring in the last exon or the 3′-most 50bp of the penultimate exon, as applied in ACMG/AMP [21].

Refined variant interpretation after RNA-seq

Variant pathogenicity was reinterpreted according to ClinGen SVI Splicing Subgroup recommendations [16] based on RNA-seq results. Determination of the PVS1(RNA) evidence code and its strength relied on: (i) the proportion of the aberrant transcript(s), derived from FRASER2 $\Delta\Psi$ or IGV inspection and accounting for allelic balance; (ii) the predicted protein consequence (e.g., frameshift, truncation) based on the observed aberrant mRNA sequence(s). These two elements determined the applicability and assigned strength (Strong, Moderate, or Supporting) of the PVS1(RNA) code.

Diagnostic outcome alteration

A change in diagnostic outcome was defined: "yes"—if the VUS was reclassified as likely pathogenic or pathogenic in a gene that would cause the patient's phenotype and the inheritance was compatible; "no"—if the pathogenicity of the variant remained uncertain.

Analysis for cases in "No candidate variant" cohort

An RNA-driven approach was utilized for this cohort. We identified significant AS events ($|\Delta\psi| \geq 0.2$, padj < 0.05) or AE events ($p < 0.05$) occurring within clinically relevant genes. The corresponding DNA sequencing data (ES/GS) for these specific genes was then inspected to identify potential underlying variants. These variants were subsequently interpreted using standard diagnostic practices (ACMG/AMP guidelines).

Number and rank of aberrant events

The number of qualifying AS/AE events was counted by applying significance thresholds (AS: $|\Delta\psi| \geq 0.2$, padj < 0.05; AE: $p < 0.05$; consistent with the RNA-driven approach in "No candidate variant" cohort) and then filtering for those involving OMIM morbid genes. The rank of a target event was its position within the filtered list of respective samples when sorted by ascending p-value.

Results

Validation of the analysis pipeline and filtering thresholds with known RNA-altering variants

To assess the effectiveness of the analysis pipeline and the filtering thresholds for AS or AE events, a set of samples harboring clinically relevant variants deemed to alter RNA splicing or expression were utilized (Supplementary Table 2). At the thresholds of $|\Delta\psi| \geq 0.2$, padj < 0.05 for AS events, AS associated with all three splicing variants were successfully identified using our outlier-based methods and annotation pipeline. At the $p < 0.05$ threshold for AE events, our pipeline successfully identified the AE events associated with: (i) a heterozygous NMD-triggering variant—a small indel in the *SON* gene—with p-value of 4.37E-06 (FC = 0.73, z = −4.77); (ii) a heterozygous deletion in the 16p13.11 region, with p-value of 1.24E-06 and 0.003 for the two OMIM morbid genes within this region—*ABCC1* (FC = 0.56, z = −5.25) and *NDE1* (FC = 0.62, z = −3.1), respectively; (iii) a triplication (four copies) of the 15q11.2 region, with p-value of 2.42E-20 and 0.0002 for two OMIM morbid genes within

Table 1 Comparison of SpliceAI prediction and RNAseq result and impact on pathogenicity interpretation and diagnostic outcome in the splicing-VUS cohort

Sample	Gene, variant and zygosity	Aberrance predicted by SpliceAI and scores (type|Δscore|pre-mRNA pos)[#]	Splicing alteration by RNA-seq	Consequence match between SpliceAI and RNA-seq	AS and AE detected by RNA-seq	Pathogenicity interpretation after RNA-seq and evidence codes	Solved with diagnostic variants?	Expression in relation to NMD, and expression ranking across all samples (ascending)	Disease and inheritance
R003	NM_007059.4 *(KPTN)*: c.863G > A, hom	Yes AL|0.43|75bp; DL|0.86|0bp;	intron 9 retention r.[863g > a; 863_864ins863 + 1_864-1], p.(Arg288Glnfs*2)	**Not match**: SpliceAI predicted exon 9 skipping	**AS**: Δψ = -0.91, p = 2.64E-17 **AE**: FC = 0.74, z = -1.95, p = 0.059	LP PVS1_ Strong(RNA) + PM2_ Supporting + PM3_ Supporting	Yes	NMD triggering, unaltered expression Rank 3rd in 128	Intellectual developmental disorder, autosomal recessive 41, AR
R044	NM_172250 *(MMAA)*: c.-66 + 2T > C, hom	Yes DL|0.90|-2bp; DG|0.43|240bp	no mRNA detected	**Not match**: SpliceAI predicted non-coding exon 1 elongation (242nt)	AS: NA **AE**: FC = 0.01, z = -11.7, p = 8.64E-32	P PVS1(RNA) + PM2_ Supporting + PM3_ Supporting + PP4	Yes	NA (variant within 5′ UTR) Rank 1st in 128	Methylmalonic aciduria, AR
R072	NM_001170629.2 *(CHD8)*: c.2730 + 4A > G, de novo het	Yes AL|0.33|247bp; DL|0.36|4bp;	intron 12 retention r.2486_2487ins2486 + 1_2487-1, p.(Gln830*); exon 13 skipping r.2487_2730del, p.(Gln830Aspfs*13)	**Partial match**: SpliceAI predicted E13 skipping	**AS**: Δψ = -0.5, p = 0.0002 AE: FC = 0.95, z = -0.68, p = 0.48	P PVS1(RNA) + PS2_ Moderate + PM2_ Supporting	Yes	NMD triggering, unaltered expression Rank 29th in 128	Intellectual developmental disorder with autism and macrocephaly, AD
R087	NM_002641.4 *(PIGA)*: c.-63 + 1G > A, maternal hemi	Yes DL|0.99|1bp; DG|0.36|-145bp	alternative donor site usage r.-63_-62ins-63 + 1_-63 + 146, p.? intron retention or alternative TSS r.?, p.?	**Partial match**: SpliceAI predicted exon 1 elongation (146nt)	**AS**: Δψ = -0.91, p = 0.0002 AE: FC = 1.04, z = 0.49, p = 0.63	VUS PM2_Sopporting	No	NA (variant within 5′ UTR) Rank 89th in 128	Multiple congenital anomalies-hypotonia-seizures syndrome 2, XLR
R095	NM_007055.4 *(POLR3A)*: c.3429 + 51A > G, maternal het, in trans with c.1771-7C > G	Yes DL|0.12|51bp; DG|0.97|1bp	alternative donor splice site usage r.3429_3430ins3429 + 1_3429 + 50, p.(Asn1145Serfs*17)	**Perfect match**: SpliceAI predicted exon 26 elongation (50nt)	**AS**: Δψ = -0.49, p = 5.40E-05 AE: FC = 0.77, z = -3.2, p = 0.003	P PVS1(RNA) + PM2_ Supporting + PM3 + PP1	Yes	NMD triggering, reduced expression Rank 1st in 128	Leukodystrophy, hypomyelinating, AR
R097	NM_001099922 *(ALG13)*: c.383 + 1067G > A, maternal hemi	Yes AG|0.20|-33bp; DG|0.15|4bp	38bp cryptic exon r.383_384ins[383 + 1034_383 + 1066;a;383 + 1068_383 + 1071], p.(Arg128Serfs*7)	**Partial match**: SpliceAI predicted the 38nt psudo-exon, but the score is low	**AS**: Δψ = -0.66, p = 1.31E-12 AE: FC = 1.06, z = 0.89, p = 0.36	VUS PM2_Sopporting	No	NMD triggering, unaltered expression Rank 104th in 128	Developmental and epileptic encephalopathy 36, XL

Table 1 (continued)

Sample	Gene, variant and zygosity	Aberrance predicted by SpliceAI and scores (type|Δscore|pre-mRNA pos)#	Splicing alteration by RNA-seq	Consequence match between SpliceAI and RNA-seq	AS and AE detected by RNA-seq	Pathogenicity interpretation after RNA-seq and evidence codes	Solved with diagnostic variants?	Expression in relation to NMD, and expression ranking across all samples (ascending)	Disease and inheritance
R111	NM_000292.3 *(PHKA2)*: c.2597 + 6T > C, maternal hemi	Yes DG|0.46|-21bp	alternative donor splice site usage r.2597_2598ins[2597 + 1_2597 + 5;c;2597 + 7_2597 + 27], p.(Pro867*)	**Perfect match**: SpliceAI predicted exon 23 elongation (27nt)	**AS**: Δψ = -0.31, p = 5.71E-08 AE: FC = 0.96, z = -0.57, p = 0.56	VUS PM2_Sopporting	No	NMD triggering, unaltered expression Rank 36th in 128	Glycogen storage disease, type IXa, XLR
R142	NM_004092.3 *(ECHS1)*: c.807 + 5G > A, maternal het, in trans with c.463G > A	Yes AL|0.12|72bp; DL|0.98|5bp; DG|0.19|-80bp	alternative donor splice site usage r.767_807del, p.(Ser256Argfs*2)	**Not match**: SpliceAI did not predicte the novel donor at 46bp upstream of the variant	**AS**: Δψ = -0.47, p = 1.95E-16 **AE**: FC = 0.72, z = -3.79, p = 0.0002	LP PVS1_Strong (RNA) + PM2_Supporting + PM3 + PP1	Yes	NMD escape, reduced expression Rank 1st in 128	Mitochondrial short-chain enoyl-CoA hydratase 1 deficiency, AR
R146	NM_007055.4 *(POLR3A)*: c.44 + 5G > C, maternal het, in trans with c.1771-6C > G	Yes DL|0.44|5bp; DG|0.62|-5bp	alternative donor splice site usage with potential novel start codon r.44_45ins[44 + 1_44 + 4;c;44 + 6_44 + 10], p.(Ile15Metfs*2)	**Perfect match**: SpliceAI predicted exon 1 elongation (10nt)	**AS**: Δψ = -0.34, p = 0.015 AE: FC = 0.98, z = -0.24, p = 0.79	VUS PM2_Sopporting	No	NA (variant within intron 1) Rank 48th in 128	Leukodystrophy, hypomyelinating, AR
R155	NM_001039591.3 *(USP9X)*: c.7432-15A > G, de novo hemi	Yes AG|0.85|1bp	alternative acceptor splice site usage r.7431_7432ins7432-14_7432-1, p.(Glu2478Ilefs*54)	**Perfect match**: SpliceAI predicted exon 44 elongation (14nt)	**AS**: Δψ = -0.76, p = 0.79 AE: FC = 1.01, z = 0.25, p = 0.80	LP PVS1_Strong (RNA) + PS2_Moderate + PM2_Supporting	Yes	NMD escape, unaltered expression Rank 78th in 128	Intellectual developmental disorder, X-linked 99, XLD/XLR

AD autosomal dominant, *AE* aberrant expression, *AG* acceptor gain, *AL* acceptor loss, *AR* autosomal recessive, *AS* aberrant splicing, *DG* donor gain, *DL* donor loss, *FC* fold change, *hemi* hemizygous, *het* heterozygous, *hom* homozygous, *LP* likely pathogenic, *NA* not available, *P* pathogenic, *VUS* variant of uncertain significance, *XLD* X-linked dominant, *XLR* X-linked recessive

this region—*UBE3A* (FC = 2.01, z = 8.25) and *HERC2* (FC = 2.01, z = 3.34), respectively, and fold change almost exactly matched the copy number; and (iv) a heterozygous deletion on the X chromosome involving *CASK*, with p-value of 0.009 (FC = 0.74), consistent with reduced expression of the mutant allele on the sex chromosome.

Diagnostic value of RNA-seq in cases with splicing VUS

Comparison between SpliceAI prediction and empirical observation in RNA-seq

The 10 candidate VUS in the test cohort were all predicted to be splice-altering based on SpliceAI (Table 1, Δscore ≥ 0.2). These variants were located in the 5' UTR exon or intron 1 splice region (n = 3, potentially altering transcript start), coding region (n = 1, with missense effect), deep intron (n = 1, over 1kb from the exon boundary), splice region (n = 4, within 20bp of exon boundaries), and intron region near splice sites (n = 1, 51bp from the exon boundary). The location of the variants and the "consequence match" between predicted and empirical RNA production are shown in Fig. 2A, B.

The predicted consequences of 40% of the variants (R095, R111, R146, R155) match perfectly with RNA-seq, 30% (R072, R087, R097) match partially and the remaining 30% (R003, R044, R142) do not match at all (Fig. 2A, B). Details of the consequences predicted by SpliceAI and observed by RNA-seq were summarized in Table 1. "Perfect match" is exemplified by the variant c.3429 + 51A > G (*POLR3A*) in R095—the prediction is consistent with the empirical observation of a 50nt elongation of exon 26 (Supplementary Fig. 1). "Partial match" is exemplified by the variant c.2730 + 4A > G (*CHD8*) in R072—while exon 13 skipping is predicted by SpliceAI, intron 12 retention was also observed in RNA-seq (Fig. 2C). "Not match" is exemplified by the

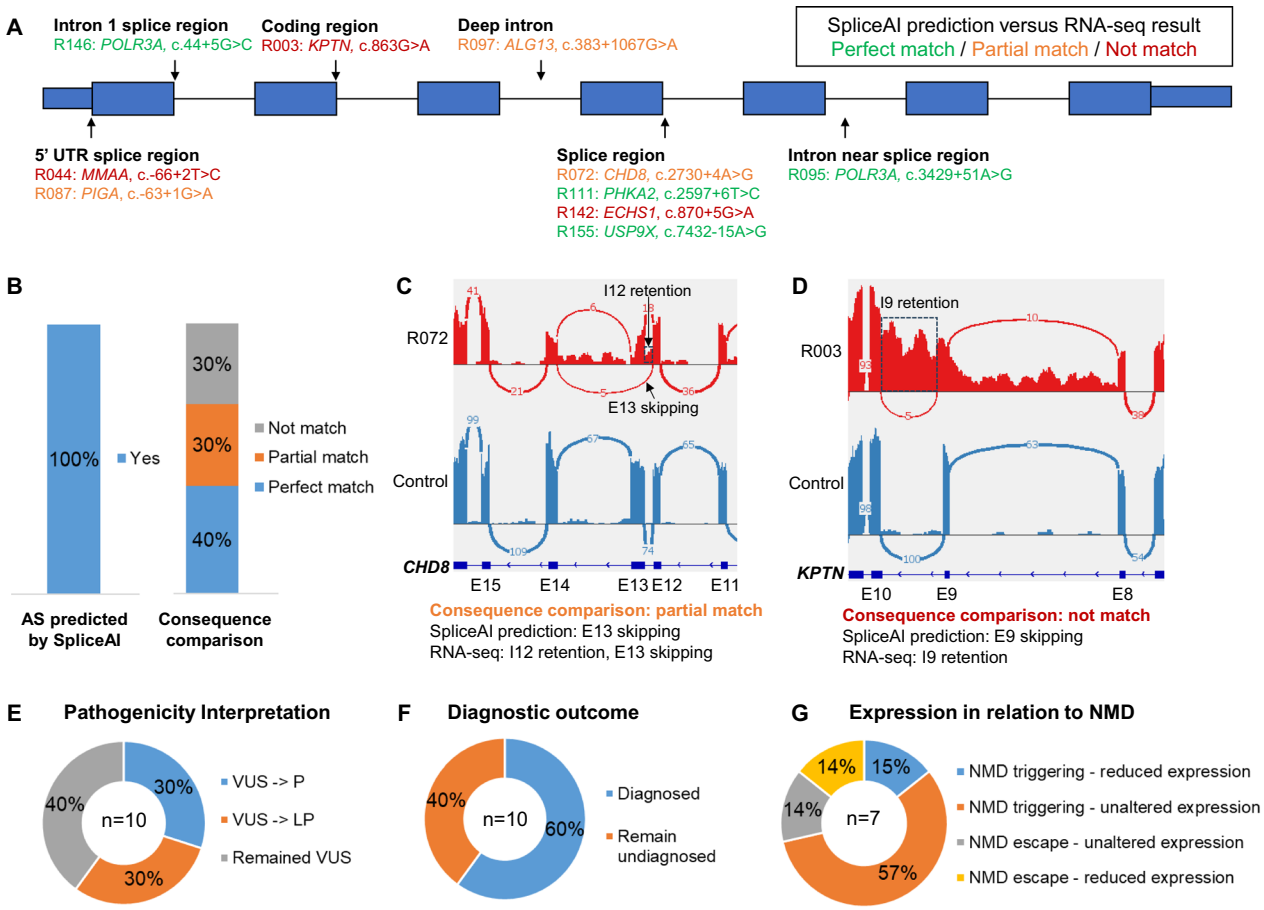

Fig. 2 Value of RNA-seq on unsolved cases with VUS. **A** The location of variants and comparison of predictions with RNA-seq results. **B** The percentage of splice donor and/or acceptor alterations predicted by SpliceAI and the degree of concordance between SpliceAI prediction and RNA-seq. **C**, **D** Sashimi plot of RNA-seq showing consequence partial or not match between SpliceAI prediction and RNA-seq. **E** The alteration of pathogenicity interpretation for candidate VUS. **F** The alteration of diagnostic outcomes for cases with candidate splicing VUS. **G** mRNA expression in relation to NMD

variant c.863G > A (*KPTN*) in R003—with this coding variant at the exon boundary, exon 9 skipping was predicted but not supported by any RNA-seq reads, instead intron 9 retention was observed (Fig. 2D).

RNA-seq altered variant interpretation and diagnostic outcome

With the empirical evidence from RNA-seq experiment, 60% of VUS were reclassified (of the ten candidate VUS, three were reclassified as pathogenic, three were reclassified as likely pathogenic, and the other four remained as uncertain significance (Fig. 2E, Table 1) according to the guideline by the ClinGen SVI Splicing Subgroup [16]. This reclassification of variants eventually changed the diagnostic outcome—molecular diagnosis was achieved in all six cases with reclassified variants, resulting in a 60% diagnostic uplift (Fig. 2F, Table 1). Two examples are described below:

VUS reclassified as pathogenic: case diagnosed R072 was a 14-month-old boy with macrosomia, hypotonia, fine motor/gross motor/language developmental delay, intellectual disability. A de novo heterozygous variant c.2730 + 4A > G of *CHD8* (associated with intellectual developmental disorder with autism and macrocephaly, autosomal dominant, OMIM#615,032) was identified. Significant AS of r.2486_2487ins2486 + 1_2487-1 (intron 12 retention) and r.2487_2730del (exon 13 skipping) and were detected by RNA-seq ($\Delta\psi = -0.5$, p = 0.0002, Table 1, Fig. 2C, Supplementary Fig. 1). With almost half reduction (based on $\Delta\psi$) of the normal junctions, the proportion of altered transcripts associated with this heterozygous variant allele was considered complete. These altered transcripts correspond to the protein alterations p.(Gln830*) and p.(Gln830Aspfs*13), respectively. Both alterations are presumed to trigger NMD, and expression analysis did show a mild but significant reduction (FC = 0.85, p = 0.003), consistent with haplo-insufficiency mechanism of *CHD8*-related disease etiology. The variant was reclassified as pathogenic with evidence codes PVS1(RNA) + PS2_Moderate + PM2_Supporting.

VUS reclassified as likely pathogenic: case diagnosed R003 was a five-year-old girl with mental retardation, speech impairment and diabetes mellitus type 1. A homozygous missense variant c.863G > A on *KPTN* (associated with intellectual developmental disorder, autosomal recessive 41, OMIM#615,637) was identified. A significant AS of intron 9 retention r.[863g > a; 863_864ins863 + 1_864-1] was detected by RNA-seq ($\Delta\psi = -0.91$, considered near completely aberrant for bi-alleles, p = 2.64E-17), corresponding to the protein alteration p.(Arg288Glnfs*2) (Table 1, Fig. 2D, Supplementary Fig. 1). The variant was reclassified as likely pathogenic with evidence codes PVS1_Strong (RNA) + PM2_Supporting + PM3_Supporting. No candidate variant associated with diabetes mellitus was identified.

RNA expression in relation to NMD triggering or escape

Based on the mis-splicing revealed by RNA-seq, seven variants in the candidate VUS pool were associated with altered splicing leading to PTC. Of these seven variants, five were presumed to trigger NMD (R003, R072, R095, R097, R111), and two were presumed to escape NMD (R111 and R155, PTC located in the last exon). Whether potential NMD-triggering or escape events could be reflected by expression alteration was assessed. In more than half cases (72%, representing 5/7 cases), the observed expression of corresponding gene was not straightforwardly consistent with predicted NMD status. Specifically, four cases harboring variants presumed to trigger NMD lacked statistically significant AE, while one case with a variant presumed to escape NMD showed significant AE in the respective gene. (Table 1, Fig. 2G), suggesting AE may not be a reliable metric for NMD status.

Diagnostic value of RNA-seq in cases without candidate variants

Diagnostic yield in the "No candidate variant" cohort

In the test cohort of 111 cases without candidate variants, three cases were eventually diagnosed with pathogenic variants associated with AE or AS events identified by RNA-seq, including two cases with AS and one case with AE (Table 2), resulting in a 2.7% increase in diagnosis in this cohort.

R020 was an eight-year-old girl with mental retardation, abnormal gait, prominent forehead, hepatosplenomegaly and multiple other anomalies. A significant AS of exon 3–4 skipping was identified in *HGSNAT* (associated with Mucopolysaccharidosis type IIIC, AR, OMIM #252,930) by RNA-seq ($\Delta\psi = -0.57$, p = 2.86E-08, padj = 0.009) (Fig. 3A). The altered mRNA of *HGSNAT* was r.235_493del, corresponding to the protein production p.(Cys79Leufs*2). The exon 3–4 deletion in trans with a missense was further validated by GS. The reason for the prior ES-negative result was due to a failed CNV calling of this 4.3kb deletion based on ES data.

R022 was an eight-year-old girl with mental retardation, speech developmental delay, triangular face, low ear position, and feeding difficulties in early childhood. AS with multiple abnormal junctions was identified in *DDX3X* (associated with Intellectual developmental disorder, X-linked syndromic, Snijders Blok type, XLD/XLR, OMIM #300958) by RNA-seq ($\Delta\psi = -0.22$,

Table 2 Additional diagnostics by RNA-seq in the "no candidate variant" cohort

Sample	DNA variant	Gene	RNA and protein product description	AE or AS detected by RNA-seq	Reasons of ES/GS-negative	Parental origin and zygosity	Disease and Inheritance	Validation of DNA
R020	seq[GRCh37] 8p11.21 (43,010,865_43,015,209) × 1	NM_152419.3 (*HGSNAT*)	r.235_493del, p.(Cys79Leufs*2)	**AS**: Δψ = -0.57, p = 2.86E-08, padj = 0.009	ES-negative, failed CNV calling for this 4.3 kb deletion	maternal, het, in trans with c.739A > G	Mucopolysaccharidosis type IIIC, AR	GS
R022	NM_001356.5 *(DDX3X)*: c.1616-12_1621del	NM_001356.5 (*DDX3X*)	r.1616_1769del, p.(Gly539Valfs*79); r.1616_1687del, p.(Gly539_Glu562del); r.1616_1621del ins1616-71_1616-13, p.(Leu540Lysfs*2)	**AS**: Δψ = -0.22, p = 1.88E-14, padj = 1.39E-08	ES-negative, poor coverage and failed variant calling	de novo, het	Intellectual developmental disorder, X-linked syndromic, Snijders Blok type, XLD/XLR	GS
R126	ogm[GRCh37] 2q23.1 (148,724,343_148,858,380) × 1	NM_018328.5 (*MBD5*)	reduction of mRNA expression	**AE**: FC = 0.42, z = -7.02, p = 2.52E-10, padj = 4.09E-05	ES-negative, no probes in 5'UTR region	de novo, het	Intellectual developmental disorder, autosomal dominant 1, AD	OGM

AD autosomal dominant, *AE* aberrant expression, *AR* autosomal recessive, *AS* aberrant splicing, *FC* fold change, *GS* genome sequencing, *het* heterozygous, *OGM* optical genome mapping, *XLD* X-linked dominant, *XLR* X-linked recessive

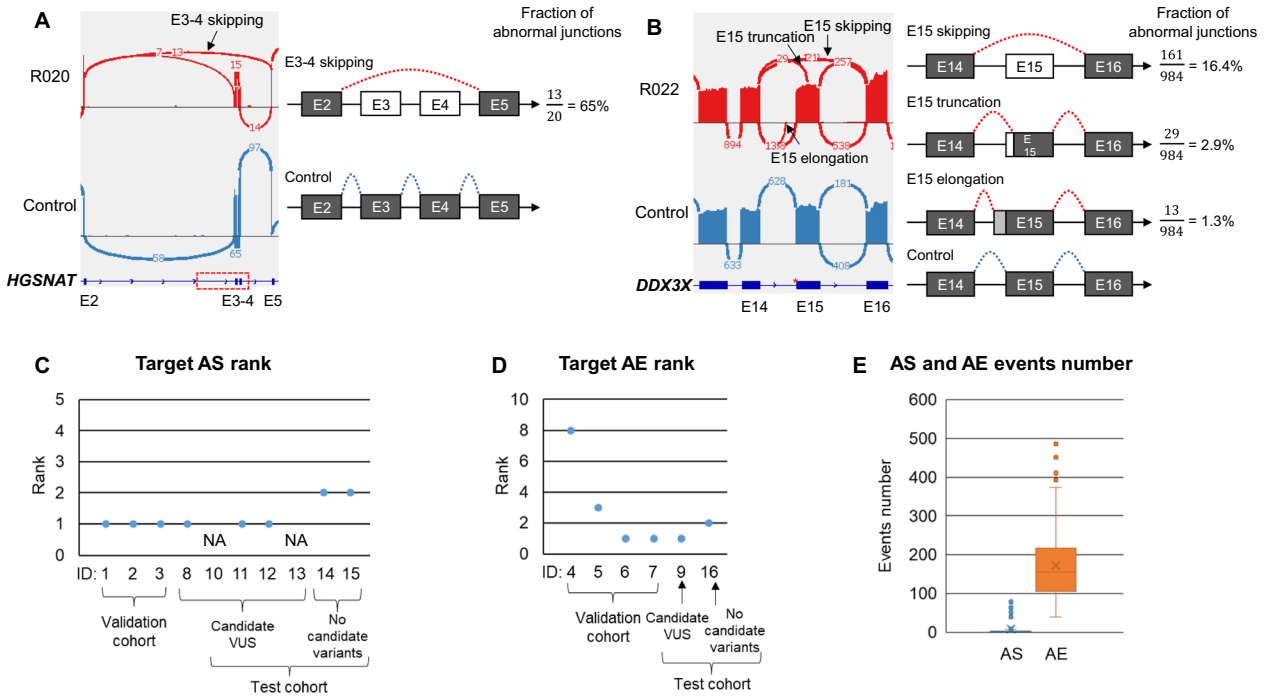

Fig. 3 Additive value of RNA-seq on unsolved cases with negative result from ES. **A**, **B** Left: Sashimi plot of RNA-seq showing abnormal junctions. Right: Pattern plot of exon junctions and the fraction of abnormal junctions. The exons 3–4 deletion in *HGSNAT* of patient R020 is indicated by a red dotted box, and the indel variant in *DDX3X* of patient R022 is indicated by a red asterisk. **C** The ranking of target AS events in FRASER ($|\Delta\psi| \geq 0.2$, padj < 0.05) for cases diagnosed with variants affecting splicing. **D** The ranking of target AE events in OUTRIDER ($p < 0.05$) for cases diagnosed with variants affecting gene expression. **E** Number of AS ($|\Delta\psi| \geq 0.2$, padj < 0.05) and AE ($p < 0.05$) events in OMIM morbid genes for the entire cohort of 128 individuals

padj = 1.39E-08). The abnormal junctions potentially lead to exon 15 skipping (16.4%), truncation (2.9%), and elongation (1.3%) (Fig. 3B). The alteration of *DDX3X* mRNA include r.1616_1769del, r.1616_1687del, and r.1616_1621delins1616-71_1616-13, corresponding to the protein alteration p.(Gly539Valfs*79), p.(Gly539_Glu562del), and p.(Leu540Lysfs*2). The proportion of DDX3X transcripts containing the aberrant splice junction is approximately 20% (Fig. 3B), which is lower than the expected 50% for a heterozygous variant and is consistent with skewed inactivation of the X chromosome. The de novo small deletion c.1616-12_1621del was further validated by GS. The reason for the prior ES-negative result was due to poor coverage of the ES in this exon boundary and failed variant calling of this 15bp deletion.

R126 was a six-year-old boy with intellectual disability, language and motor developmental delay. Significantly decreased mRNA expression of *MBD5* (associated with Intellectual developmental disorder, X-linked 99, XLD/XLR, OMIM #156200) was identified by RNA-seq (FC = 0.42, z = − 7.02, p = 2.52E−10, padj = 4.09E−05). The de novo deletion of a non-coding region affecting the promoter and 5' UTR was further validated by chromosomal microarray analysis. The reason for the prior ES-negative result was due to lack of coverage in this non-coding region. This case has been described in detail with implications from optical genome mapping in another study [11].

Potential and feasibility of an RNA-driven approach regardless of pre-existing information on candidate variants

One question that remains for clinical practice is whether an RNA-driven approach is effective and feasible, regardless of pre-existing information on candidate variants. To assess the potential and feasibility of the RNA-driven approach, the ranking of target AS and AE events in final diagnosed cases (both in the validation cohort and in the test cohort with or without candidate variants) was evaluated based on our analysis pipeline (see Methods for detailed ranking algorithm and Supplementary Table 4). In the 10 cases diagnosed with splice-altering variants (Fig. 3C), the target AS events were ranked in the top 2 in 8 cases (80%, including 6 cases with a top 1 ranking), but were not called in the other two cases (20%, padj did not pass the filtering threshold). In the 6 cases diagnosed with expression-altering variants, the target AE events were ranked in the top 3 in 5 cases (83%) and top 8 in all cases (Fig. 3D).

The number of potential candidates for further reviewing based on the RNA-driven approach was also assessed. For aberrations passing the filtering thresholds and involving OMIM morbid genes, the median number of candidate AS events was 1 (range: 0–85, Q1 = 0, Q3 = 3), and the median number of candidate AE events for OMIM morbid genes was 118.5 (range: 0–476, Q1 = 85, Q3 = 161). The distribution of the number of AS and AE events in OMIM morbid genes for the entire cohort of 128 individuals is shown in Fig. 3E.

Discussion

Diagnostic uplift by RNA-seq

In this study, we evaluated the additive value of RNA-seq in ES/GS unsolved cases and found a 60% diagnostic yield in cases with splicing VUS (6/10) and a 2.7% yield in cases without candidate variants (3/111), culminating in an overall yield of 7.4% (9/121) for the entire cohort. The diagnostic rate in the test cohort with candidate splicing VUS is similar to previous studies using RNA diagnostics in clinically accessible samples, which reclassified 75% of putative splicing variants [10]. The overall diagnostic yield of RNA analysis in previous studies was 7.5–34% [5–8, 13, 14, 22]. Variations in diagnostic yields across studies can be attributed to differences in cohort composition, particularly the proportion of cases with pre-existing uncertain splicing variants or indicated genes, as well as the source tissues used for analysis. For instance, a recent study utilizing blood RNA-seq in a similar setting reported a diagnostic rate of 7.3% (4 out of 55 cases) in individuals without pre-existing candidates, compared to a diagnostic rate of 35% (18 out of 51 cases) in those with splicing VUS [13]. While reporting practices vary, these findings highlight that cohort composition remains a key determinant of the overall diagnostic yield reported in RNA-seq studies.

RNA analysis refines interpretation of candidate splice variants

Most splice variants outside the classic donor/acceptor sites have been classified as VUS based on DNA-level evidence alone. The precise nature of the mRNA product—the exact sequence, the presence or absence of a PTC, and the specific coding regions affected—is essential for determining pathogenicity. While SpliceAI is superior at predicting whether a variant alters splicing (a binary yes or no prediction) [15], it does not accurately predict the final mRNA product. Furthermore, the proportion of altered transcripts, a critical factor in current interpretation guidelines (e.g. distinguishing between complete and incomplete mis-splicing) [10], is not provided by SpliceAI. The proportion of aberrant transcripts may also implicate disease severity. For example, a hemizygous variant of *USP9X* that only partially affects splicing in a male patient (R155) could explain the patient's relatively moderate phenotypes, similar to those observed in female patients [23].

Our results revealed a critical limitation of current DNA-based prediction: approximately 60% of SpliceAI-predicted alterations did not perfectly match RNA-seq-observed outcomes (i.e., a partial match or a complete mismatch). Variant classification factors in the functional impact of the predicted mRNA consequence—specifically the strength of evidence code PS1 and PVS1 hinges on the functional impact—which could ultimately influence diagnostic conclusions [16]. In our study, the discrepancies between SpliceAI and RNA-seq outcome ("partial match" and "mismatch") primarily arose from inaccurate predictions of the combined effects on acceptor and donor sites, leading to differences in observed outcomes, such as exon skipping versus intron retention.

In clinical practice, reverse transcription polymerase chain reaction (RT-PCR) has traditionally been used to validate splicing events when a specific alteration is suspected [10, 12]. RT-PCR is highly targeted, cost-effective, and feasible for genes with low expression in accessible tissues. However, its utility is inherently limited by its dependence on prior knowledge of potential splicing events for primer design. In contrast, RNA-seq offers a hypothesis-free approach that captures the full spectrum of splicing events without primer bias. Given our observations of discordance between computational predictions (e.g., SpliceAI) and observed RNA outcomes, RNA-seq should be advocated as the preferred strategy whenever gene expression is sufficient to support its use, thereby improving the accuracy of splicing characterization and ultimately improving diagnostic outcomes.

Observed expression in relation to NMD

NMD is a surveillance system that degrades mRNAs containing a PTC and plays an important role in protein homeostasis and disease. The expression levels observed in this study do not correlate perfectly with the presumed NMD trigger or escape (according to the rule that NMD would occur unless the PTC is located in the first exon, the last exon or in the last 50bp of the penultimate exon) [21]. Variability in NMD efficacy may arise from factors such as the exon sequence downstream of the PTC, the PTC-to-intron distance, and the number of introns flanking the PTC [24]. According to the previous study providing genome-wide predictions of NMD efficacy, NMD efficacy is highly variable across genes: only in 17% of human genes, > 75% of the coding regions allow NMD to be fully triggered when a PTC occurs; whereas in 36% of genes, > 75% of the coding sequence allows partial NMD evasion [25, 26]. Additionally, NMD activity observed

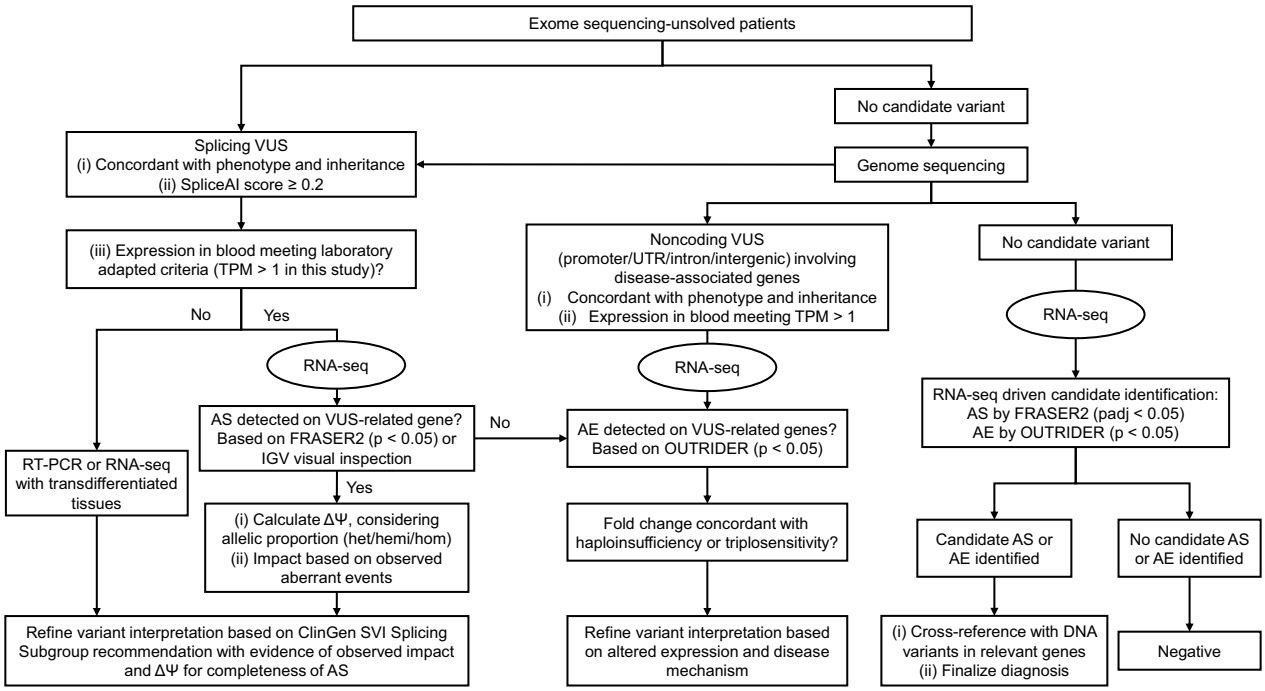

Fig. 4 The proposed workflow for conducting RNA-seq and integrating the results into variant curation process

in blood RNA may not fully capture the tissue-specific regulatory dynamics of mRNA surveillance. Therefore, when FRASER2 or IGV unambiguously reveals an AS event, we rely solely on PTC location-dependent NMD prediction and do not consider AE data—because NMD efficiency and tissue specificity can vary, making AE an unreliable triage metric in this context. Conversely, efficient NMD can eliminate abnormal transcripts, preventing AS outliers from being detected. If no splice aberration is observed, we then examine AE and interpret the fold-change in expression in light of the known disease mechanism (e.g., haploinsufficiency). The workflow for variant reinterpretation based on RNA-seq results is shown (Fig. 4).

RNA-driven versus RNA-complementary approach: which is preferred strategy?

The preferred timing of RNA-seq test relative to ES, GS, and other genomic technologies (e.g., optical genome mapping, long-read sequencing) remains an open question. While the benefits of RNA analysis are widely recognized, the key decision lies in the choice between an RNA-driven and an RNA-complementary approach. An RNA-driven approach prioritizes the identification of candidate genes by RNA-seq, followed by targeted testing at the DNA level to investigate variants within the candidate genes. The feasibility of this strategy depends on the ability of RNA-seq to identify a manageable number of candidate genes or regions with sufficient sensitivity. Conversely, an RNA-complementary approach typically involves RNA-seq in conjunction with or following genome-scale DNA testing, primarily to interpret the functional consequences of variants identified by DNA-based methods.

To evaluate the feasibility of an RNA-driven approach, we assessed the ranking of AS and AE events detected by RNA-seq in all diagnosed cases within our cohort. In the three cases solved without pre-existing candidate VUS (the "No candidate variant" cohort), the causative AS or AE events were among the top three ranked outliers. Similarly, in the validation cohort, the aberrant events associated with the known pathogenic variants were ranked in the top three in six out of seven cases (with one sample, R008, ranked eighth; Supplementary Table 4). A critical measure of the RNA-driven approach is its ability to identify cases with VUSs without relying on prior candidate VUS variants. This measure could implicate the potential utility of RNA-seq before comprehensive genome-scale testing (e.g., WGS, optical genome mapping, or long-read sequencing). Our results showed that four out of six cases with pre-existing candidate variants could have been identified based on RNA-driven approach. However, two cases would have been missed (33%), suggesting that an RNA-driven approach, while valuable, may not be fully comprehensive.

In our study, the three cases (R020, R022, R126) resolved by RNA-seq in the "No candidate variant" cohort were all ES-negative due to lack of coverage or failed variant calling (see Table 2, "Reasons for negative finding in prior testing"). While performance can vary between ES capture and pipelines, the inherent difficulties in reliably detecting small CNVs and variants in poorly covered noncoding regions remain universal challenges for exome sequencing. However, all these three cases could have been solved by GS, which provides more uniform coverage and includes non-coding regions [27]. No additional diagnoses were made by RNA-seq in the GS-negative cases. Three cases (R072, R095, R142) with VUS identified by GS (all intron variants) were reclassified as pathogenic or likely pathogenic based on RNA-seq data, leading to a definitive molecular diagnosis. In this regard, the RNA-complementary approach would be favored within the current diagnostic paradigm. The proposed workflow for conducting RNA-seq and integrating the results into variant curation process is shown in Fig. 4.

Limitations of this study

Whole blood is one of the most accessible tissues in the clinical setting, but a significant number of OMIM genes have low expression levels in blood. Blood expression of OMIM genes varies considerably across disease categories, potentially introducing bias into studies such as ours, depending on the specific cases referred to. Advances in the transdifferentiation approach hold great promise for future RNA diagnostics using clinically accessible tissues [6].

Another limitation is the lack of a widely accepted pipeline and metrics for RNA-seq analysis in diagnostic settings. Our study used the commonly used DROP pipeline, but without validation against other algorithms, which may limit the comprehensiveness of the results. Furthermore, this study prioritizes clinical utility, focusing on its relevance to diagnostic outcomes, rather than providing a comprehensive assessment of analytical performance. Consequently, validation with a larger cohort or extensive optimization of pipelines and thresholds are beyond the scope of this work. The thresholds applied were derived from published studies and adjusted or confirmed empirically with our validation cohort. Specifically, the cutoff for AE analysis – FDR corrected padj < 0.05 was used in literature [7]. However, this may be overly stringent, failing to identify 4 out of 5 known positive samples in our validation cohort (details in Supplementary Table 2). We therefore employed a $p < 0.05$ threshold which successfully recovered all true positives in our validation set. This choice of p-value is aligned by a recent study which demonstrated substantially higher sensitivity of $p < 0.05$ for AE detection compared to FDR corrected thresholds [28]. We acknowledge this threshold selection could influence the overall results with relaxed statistical stringency. For readers interested in a deeper exploration of analytical considerations pertinent to clinical RNA-seq validation, we refer them to a recent study that comprehensively addresses these aspects [28].

Conclusions

These findings underscore the utility of blood RNA-seq to improve diagnostic outcomes in an ES/GS-unsolved rare disease cohort. Our results demonstrate high efficacy of RNA-seq in the presence of splicing VUS; although solving "No candidate variant" cases via an RNA-driven approach is feasible, it remains inherently challenging. We therefore advocate the use of RNA-seq to complement ES/GS investigations in rare disease diagnostics.

Abbreviations

AE	Aberrant expression
AS	Aberrant splicing
CNV	Copy number variation
ES	Exome sequencing
FC	Fold change
GS	Genome sequencing
NMD	Nonsense-mediated mRNA decay
PTC	Premature termination codon
RNA-seq	RNA sequencing
VUS	Variant of uncertain significance

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06609-w.

Supplementary Material 1. Supplementary Table 1. Basic information for the patient recruited in our cohort and total AS, AE events on OMIM morbid gene by RNA-seq. Supplementary Table 2. Variant information, aberrations detected by RNA-seq and altered production description of the validation cohort. Supplementary Table 3. Parameters used in DROP analysis. Supplementary Table 4. Rank of the target AS or AE events in RNA-seq diagnosed cases

Supplementary Material 2. Supplementary Fig. 1. Pattern plot of exon junctions and Sashimi plot of RNA-seq showing abnormal junctions

Acknowledgements

The authors thank all the patients and their family members who participated in this work.

Author contributions

The study was designed by YF and YY. Patients were recruited by YY, BX, LL, and KZ. Samples were collected by YL. Data analysis and interpretation was done by YF, XL, TX and HL. The first draft of the manuscript was written by XL and YF, and it was reviewed by YY, YF, and XL. All authors read and approved the final version of the manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (No.2022YFC2703400 to YY; No.2022YFC2703405 to YF), National Natural Science Foundation of China (No.82171165 to YF; No.82271904 to YY), Shanghai Oriental Talents (No.QNWS2024020) and Science and Technology Commission of Shanghai Municipality (No.24Y12800700 to YY).

Availability of data and materials
Information was uploaded to LOVD for the 13 patients described in Tables 1 and 2 (https://databases.lovd.nl/shared/screenings). They were recorded in the database as individuals #00464292-00464301 and #00464315-00464317. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the principles of the Declaration of Helsinki. This study was approved by the ethic board of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (XHCC-C-2022-077-1). Written informed consent for genetic analysis was obtained from the patients or their legal guardians.

Consent for publication
Consent was obtained from the patients and their legal guardians for patient-related information that could be published.

Competing interests
The authors declare no competing interests.

Received: 14 March 2025 Accepted: 13 May 2025
Published online: 26 May 2025

References

1. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med. 2013;369(16):1502–11.
2. Taylor JC, Martin HC, Lise S, Broxholme J, Cazier JB, Rimmer A, et al. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. Nat Genet. 2015;47(7):717–26.
3. 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(3):466–83.
4. Zhang B, Gao X. Deciphering DNA variant-associated aberrant splicing with the aid of RNA sequencing. Nat Genet. 2023;55(5):732–3.
5. 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(6):911–9.
6. Li S, Zhao S, Sinson JC, Bajic A, Rosenfeld JA, Neeley MB, et al. The clinical utility and diagnostic implementation of human subject cell transdifferentiation followed by RNA sequencing. Am J Hum Genet. 2024;111(5):841–62.
7. Yépez VA, Gusic M, Kopajtich R, Mertes C, Smith NH, Alston CL, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. Genome Med. 2022;14(1):38.
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. https://doi.org/10.1126/scitranslmed.aal5209.
9. 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.
10. Bournazos AM, Riley LG, Bommireddipalli S, Ades L, Akesson LS, Al-Shinnag M, et al. Standardized practices for RNA diagnostics using clinically accessible specimens reclassifies 75% of putative splicing variants. Genet Med. 2022;24(1):130–45.
11. Xiao B, Luo X, Liu Y, Ye H, Liu H, Fan Y, et al. Combining optical genome mapping and RNA-seq for structural variants detection and interpretation in unsolved neurodevelopmental disorders. Genome Med. 2024;16(1):113.
12. He WB, Xiao WJ, Dai CL, Wang YR, Li XR, Gong F, et al. RNA splicing analysis contributes to reclassifying variants of uncertain significance and improves the diagnosis of monogenic disorders. J Med Genet. 2022;59(10):1010–6.
13. Jaramillo Oquendo C, Wai HA, Rich WI, Bunyan DJ, Thomas NS, Hunt D, et al. Identification of diagnostic candidates in Mendelian disorders using an RNA sequencing-centric approach. Genome Med. 2024;16(1):110.
14. 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. https://doi.org/10.1172/JCI141500.
15. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell. 2019;176(3):535-48.e24.
16. Walker LC, Hoya M, Wiggins GAR, Lindy A, Vincent LM, Parsons MT, et al. Using the ACMG/AMP framework to capture evidence related to predicted and observed impact on splicing: Recommendations from the ClinGen SVI Splicing Subgroup. Am J Hum Genet. 2023;110(7):1046–67.
17. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
18. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–6.
19. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. Bioinformatics. 2012;28(16):2184–5.
20. 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(2):1276–96.
21. Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018;39(11):1517–24.
22. Dekker J, Schot R, Bongaerts M, de Valk WG, van Veghel-Plandsoen MM, Monfils K, et al. Web-accessible application for identifying pathogenic transcripts with RNA-seq: Increased sensitivity in diagnosis of neurodevelopmental disorders. Am J Hum Genet. 2023;110(2):251–72.
23. Johnson BV, Kumar R, Oishi S, Alexander S, Kasherman M, Vega MS, et al. Partial loss of USP9X Function leads to a male neurodevelopmental and behavioral disorder converging on transforming growth factor β signaling. Biol Psychiatry. 2020;87(2):100–12.
24. Hoek TA, Khuperkar D, Lindeboom RGH, Sonneveld S, Verhagen BMP, Boersma S, et al. Single-molecule imaging uncovers rules governing nonsense-mediated mRNA decay. Mol Cell. 2019;75(2):324-39.e11.
25. Lindeboom RGH, Vermeulen M, Lehner B, Supek F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. Nat Genet. 2019;51(11):1645–51.
26. Supek F, Lehner B, Lindeboom RGH. To NMD or Not To NMD: nonsense-mediated mRNA decay in cancer and other genetic diseases. Trends Genet. 2021;37(7):657–68.
27. Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. Genome Med. 2022;14(1):23.
28. Zhao S, Macakova K, Sinson JC, Dai H, Rosenfeld J, Zapata GE, et al. Clinical validation of RNA sequencing for Mendelian disorder diagnostics. Am J Hum Genet. 2025;112(4):779–92.

Publisher's Note