

RNA sequencing significantly improves HYDIN diagnosis and defines the effects of common deletion in primary ciliary dyskinesia

Claire Jackson

C.L. JACKSON@soton.ac.uk

University of Southampton, UK

Carolina Jaramillo Oquendo

University of Southampton, UK

Janice Coles

University of Southampton, UK

Lee Baker

University of Southampton, UK

David Bunyan

Wessex Regional Genetics Laboratory <https://orcid.org/0000-0001-8686-558X>

N. Simon Thomas

Salisbury District Hospital

James Thompson

Gemma Fryatt

Sajida Khan

David Hunt

Wessex Clinical Genetics Service, University Hospitals Southampton NHS Foundation Trust

John Holloway

University of Southampton <https://orcid.org/0000-0001-9998-0464>

Jane Lucas

University of Southampton <https://orcid.org/0000-0001-8701-9975>

Diana Baralle

Southampton General Hospital, University of Southampton <https://orcid.org/0000-0003-3217-4833>

Woolf Walker

Article

Keywords: RNA-seq, HYDIN, primary ciliary dyskinesia, diagnostics

Posted Date: May 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9545795/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is no duality of interest

Version of Record: A version of this preprint was published at European Journal of Human Genetics on October 16th, 2026. See the published version at <https://doi.org/10.1038/s41431-026-02243-2>.

Abstract

Background

Primary Ciliary Dyskinesia (PCD) is a rare genetic disorder affecting ~ 1:7,500 individuals characterised by recurrent chronic respiratory infections, bronchiectasis, laterality defects, and infertility. While at least 60 PCD-causing genes are reported, approximately 30% of patients remain without confirmed molecular diagnosis. PCD caused by biallelic *HYDIN* variants is difficult to diagnose due to a 98% homologous pseudogene (*HYDIN2*). Additionally, functional ciliary diagnostic testing can be inconclusive due to lack of discernible ciliary ultrastructure defects by electron microscopy. Our aim was to diagnose genetically unresolved PCD patients with strong clinical suspicion of *HYDIN* defects using RNA sequencing (RNA-seq).

Methods

RNA-seq was performed on total RNA extracted from patients' nasal epithelium cultured for 21-days at air-liquid interface. Splicing abnormalities were detected using rMATS and genome aligned reads manually inspected using Integrative Genome Viewer.

Results

Thirteen individuals had strong clinical suspicion of PCD, normal TEM and inconclusive genetics: 11/13 had a single pathogenic or likely pathogenic variant in *HYDIN*, 1/13 had no variants, and 1/13 required clarification of variant origin (*HYDIN* vs *HYDIN2*). RNA-seq resolved 85% of cases. The same complex pathogenic aberrant splicing events were seen in seven unrelated patients. This abnormal splice pattern was caused by a large 10 kilobase deletion spanning intron 17 to intron 18.

Conclusion

RNA-seq increases the diagnostic yield for patients with suspected PCD and has exposed a *HYDIN* 'hotspot' variant, increasing observed prevalence of *HYDIN*-related cases within our genetically diagnosed cohort. Routine PCD genomics testing will now incorporate investigation of this genetic region of *HYDIN*.

Introduction

Primary ciliary dyskinesia (PCD) is a rare genetic disorder of motile cilia estimated to affect ~ 1:7,500 individuals globally¹. It affects multiple systems, causing chronic sinopulmonary infections and bronchiectasis, laterality defects and subfertility². PCD is a heterogeneous disorder, usually inherited in

an autosomal recessive manner, with mutations in genes responsible for cilia structure and function. According to international consensus guidelines³, diagnosis of PCD can be confirmed by either transmission electron microscopy (TEM) quantification of a class 1 ciliary ultrastructure defect or identification of bi-allelic or hemizygous pathogenic or likely pathogenic variants in evidenced PCD genes⁴. Guidelines recommend adjunct tests such as high-speed video analysis (HSVA) of ciliary beat pattern and frequency, immunofluorescence (IF) localisation of ciliary proteins and measurement of nasal nitric oxide (nNO) to support highly likely diagnoses. Despite at least 60 known PCD-associated genes, ~ 30% of patients remain without a genetic diagnosis after blood genomics-panel testing; due to negative or single heterozygous results or variants of uncertain significance with non-specific ciliary function testing precluding variant classification via ACMG guidelines⁵.

PCD patients harbouring variants in the *HYDIN* gene can be particularly challenging to diagnose as ciliary ultrastructure appears normal to the human eye by TEM. *HYDIN* protein localises to the C2b subunit of the central apparatus within the ciliary axoneme (Fig. 1). Because nodal cilia lack a central pair, *HYDIN* is not required for left-right patterning, and patients with *HYDIN* variants do not present with laterality defects. There is no commercially available antibody to reliably localise *HYDIN* protein to the ciliary axoneme, however loss of immunolocalisation of the C1b projection protein SPEF2 is seen in *HYDIN* patients, and its use is advocated to assess for *HYDIN*-suspected cases in PCD diagnostic pathways⁶.

Genetic diagnosis is confounded by the presence of the 98% homologous pseudogene (*HYDIN2*) located on a different chromosome and not expressed in the lung^{7,8}. Mismapping DNA sequence reads and false identification of pathogenic variants within *HYDIN* (16q22.2) are caused by the confounding presence of *HYDIN2* (1q21.1)⁹. Bioinformatic analytical pipelines routinely filter out sequence reads that map to > 1 genomic location, leading to reduced coverage of *HYDIN*. The PCD community recognise that *HYDIN*-related PCD is underdiagnosed and in some countries excluded from genetic tests because of these difficulties.

RNA sequencing (RNA-seq) is a complementary tool to DNA testing to increase rare disease diagnoses. Within rare disorders, transcriptomic analyses have identified variants usually missed by DNA testing, often identifying non-coding variants affecting splicing that were filtered out in the initial analysis or not included in panel or whole exome sequencing testing^{10–14}. RNA-seq is an agnostic testing approach, allowing us to demonstrate differential gene expression and identify aberrant gene splicing to support a genetic diagnosis of PCD¹⁵.

Here we show that RNA sequencing markedly improves diagnostic yield in PCD patients with unresolved molecular diagnosis and a high suspicion of *HYDIN* defects following specialist ciliary function analysis within a PCD specialist centre that is adherent to the 2025 European Respiratory Society and American Thoracic Society (ERS/ATS) diagnostic guidelines³.

Material and Methods

Patient cohort

The study cohort comprises thirteen 'highly likely' or 'suspected' PCD patients, including one pair of affected siblings (P2 and P3), who had undergone PCD diagnostic clinical examination, nNO and ciliary function diagnostic testing including HSVA, IF and TEM. Participants gave written consent for the use of their samples and data in accordance with approval granted by Southampton and South West Hampshire Research Ethics Committee A (CHI395, 06/Q1702/109).

Nasal nitric oxide measurement

Nasal nitric oxide was measured during velum closure by either breath hold or oral exhalation against resistance. The highest nNO measurement was recorded from 3 reproducible tests. Tests were performed on either the NIOX Flex or Ecomedics CLD 88 Exhalyzer chemiluminescent analysers according to international standards¹⁶.

High speed video

Ciliary beat pattern was analysed on fresh nasal epithelial cells brushed from patients' inferior turbinates using a $\times 100$ objective lens, with samples equilibrated to 37°C. Videos were recorded at 500 frames per second and analysed at 30–60 frames per second from a minimum of 6 strips of ciliated epithelium. Patient cells were cultured at air-liquid interface (ALI) as previously described¹⁷, with beat pattern analysis repeated on recapitulated cilia.

Immunofluorescence labelling

Nasal epithelial cells were air dried onto slides, fixed with 4% paraformaldehyde and immunolabelled with primary antibodies anti-SPEF2 (HPA039606 Merck, Germany) 1:150 and anti-alpha tubulin (T9026 Merck, Germany) 1:1000, followed by secondary antibodies Alexafluor 488 and 594 (A21121, A11012 Life Technologies, USA) 1:2500.

Transmission Electron Microscopy

Quantification of ciliary ultrastructure was carried out on a minimum of 100 ciliary axonemes and reported according to international consensus guidelines⁴. In cases with normal cilia ultrastructure and high suspicion of PCD on beat pattern analysis, PCD-Detect-TEM, an image averaging and alignment package for investigation of ciliary features¹⁸, was used to resolve the C2b component of the central pair complex.

Genetic testing

Patient blood samples for 10 patients were sent to Wessex Genomics Laboratory Service (Salisbury) for routine PCD diagnostic genetic testing against the R189 panel (Respiratory ciliopathies including non-CF bronchiectasis (Version 4.55)). Genetic testing for the remaining three patients was carried out at the Royal Brompton Hospital.

Variant review and classification were performed by the Wessex Regional Genetics Laboratory according to ACMG/AMP guidelines and current Association for Clinical Genomic Science (ACGS) variant classification guidelines¹⁹.

The recurrent *HYDIN* exon 18 deletion was defined as c.2377 – 762_2529 + 9003delinsTATTA using primers flanking the region (for – tagcctactcaagctaattgtg, rev – aatgtgacataggcggttacc). The variant was confirmed to be present in *HYDIN* as the pseudogene differs by six base pairs across the sequenced region. The addition of an extra forward primer into the reaction from within the deleted region (tttgagacagagtctcgctc) allowed the determination of zygosity.

Short read RNA sequencing and bioinformatics pipeline

RNA was isolated from day 21 ALI-cultures by RNeasy Plus Mini Kit (Qiagen). Samples with an Agilent Bioanalyzer RIN score > 7 underwent total cDNA library preparation (NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina®) with ribosomal RNA depletion (Illumina® Ribo-Zero Magnetic Kit). Samples were sequenced by Novogene on an Illumina HiSeq2500, with 150 base pair-end reads at a sequencing depth of 70 million. Raw FASTQ files were aligned to the human genome reference (GRCh38) using STAR aligner v2.7.10a²⁰. The coordinates for *HYDIN2* were extracted into a Browser Extensible Data (BED) file. Using these coordinates, the maskfasta function from bedtools was used to mask and remove *HYDIN2* from GRCh38²¹. Control samples were previously sequenced in preliminary work at a depth of 100 million¹⁵.

rMATS-turbo was used to detect aberrant splicing across all samples and was run with additional parameters –novelSS to enable detection of novel splice sites, as well as –allow-clipping to allow alignments with soft or hard clipping to be used²². To determine the functional consequence at a transcript level for each variant, RNA-seq data was loaded into the Integrative Genomics Viewer (IGV)²³ and each variant was examined as previously detailed¹². Sashimi plots were generated with ggsashimi v.1.1.5²⁴.

Long read RNA sequencing

RNA was isolated from 21-day ALI cultures using RNeasy Plus Mini Kit (Qiagen). cDNA synthesis was conducted using the Iso-Seq Express 2.0 kit, followed by library construction using the Kinnex Full-Length RNA Kit. Samples were sequenced by Novogene on two SMRT Cells. Circular consensus (HiFi) reads generated by the sequencing provider were supplied in BAM format. These reads had undergone

standard PacBio SMRT Link processing, including circular consensus sequence generation and adapter removal. HiFi reads were aligned to the human reference genome (GRCh38) using minimap2 wrapper pbmm2 v.26.1.99 with ISOSEQ alignment mode²⁵. FLAIR was used to generate a transcriptome of high confidence isoforms with the transcriptome module²⁶.

Results

All 13 patients in the cohort had clinical characteristics consistent with PCD, with rhino-sinus and otologic disease common, and chronic wet cough and neonatal respiratory distress reported in all but proband P4. All patients had normal situs, consistent with PCD caused by *HYDIN* variants. Nasal nitric oxide was below the diagnostic cut off threshold of 77 nL/min in all but patient P2. TEM analysis provided a normal ciliary ultrastructure outcome in all patients; however, PCD-Detect-TEM averaging of TEM images showed a volume deficit in the central apparatus in nine patients, three probands yielded inconclusive results, and one showed an intact central apparatus (**Supplementary Fig. 1**). Ciliary beat pattern analysis by HSVA for all probands was abnormal with a reduced beat amplitude, lack of coordination and presence of some semi-circling and static cilia both pre- and post-ALI culture. Anti-SPEF2 localisation was reduced or absent compared to controls in 11 patients, however patient P6 demonstrated normal anti-SPEF2 IF localisation at intensity similar to controls on their pre-ALI culture (original biopsy) sample, which became absent in the post-ALI culture sample, and in patient P13 normal SPEF2 localisation was observed pre- and post-ALI culture (**Supplementary Fig. 1**).

Following PCD gene panel testing, 11 of the 13 patients had a single pathogenic or likely pathogenic variant in *HYDIN* (where bioinformatics pipeline could differentiate between *HYDIN* and *HYDIN2*) but no second variant. Of the remaining patients, one had no PCD gene variants (negative) and the remaining patient had two pathogenic variants that aligned within the 98% homologous exons present in both *HYDIN* and *HYDIN2*, meaning they could not be interpreted. A summary of the cohort is found in Table 1. Two of the cases (P1 and P9) have been previously published¹⁵.

Table 1
Cohort description.

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
P1*	10	nNO: 10 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, rotation HSVA post-ALI: static TEM: normal Detect: C2b absent IF SPEF2: absent	Heterozygous variant NM_001270974.2:c.8973delT p. (Glu2992ArgfsTer27)
P2†	4.5	nNO: 173 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: nil NRDS: nil	HSVA pre-ALI: reduced coordination and amplitude HSVA post-ALI: reduced coordination and amplitude TEM: normal Detect: C2b inconclusive IF SPEF2: reduced	Heterozygous variant NM_001270974.2:c.13402G > T p. (Val4468Phe)

Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. **nNO:** nasal nitric oxide; **NRDS:** neonatal respiratory distress syndrome; **HSVA pre-ALI:** High speed video microscopy of ciliary beat pattern before ALI culture; **HSVA post-ALI:** High speed video microscopy of ciliary beat pattern after ALI culture; **TEM:** Transmission electron microscopy of ciliary ultrastructure; **Detect:** to visualise central apparatus volume within averaged TEM images; **IF SPEF2:** Immunofluorescence labelling of SPEF2 protein.

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
P3†	0.2	nNO: 61.7 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: hearing loss NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, slight rotation in one area HSVA post-ALI: reduced coordination and amplitude TEM: normal Detect: C2b absent IF SPEF2: reduced	Heterozygous variant NM_001270974.2:c.13402G > T p. (Val4468Phe)
P4‡	29	nNO: 24.6 nL/min Cough: none reported Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear NRDS: nil	HSVA pre-ALI: reduced coordination and amplitude, some static HSVA post-ALI: static and residual movement TEM: normal Detect: C2b absent IF SPEF2: reduced/absent	Heterozygous variant NM_001270974.2:c.1095del p. (Phe365LeufsTer64)

Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. **nNO:** nasal nitric oxide; **NRDS:** neonatal respiratory distress syndrome; **HSVA pre-ALI:** High speed video microscopy of ciliary beat pattern before ALI culture; **HSVA post-ALI:** High speed video microscopy of ciliary beat pattern after ALI culture; **TEM:** Transmission electron microscopy of ciliary ultrastructure; **Detect:** to visualise central apparatus volume within averaged TEM images; **IF SPEF2:** Immunofluorescence labelling of SPEF2 protein.

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
P5	11.3	nNO: 8 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear and hearing loss NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, some static HSVA post-ALI: reduced coordination and amplitude, some static TEM: normal Detect: C2b inconclusive IF SPEF2: reduced	Panel negative
P6	5.4	nNO: 23.1 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, some static HSVA post-ALI: residual movement, some circling TEM: normal Detect: C2b absent IF SPEF2: pre-ALI present, post-ALI absent	Heterozygous variant NM_001270974.2:c.15050T > C p. (Leu5017Pro)
Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. nNO: nasal nitric oxide; NRDS: neonatal respiratory distress syndrome; HSVA pre-ALI: High speed video microscopy of ciliary beat pattern before ALI culture; HSVA post-ALI: High speed video microscopy of ciliary beat pattern after ALI culture; TEM: Transmission electron microscopy of ciliary ultrastructure; Detect: to visualise central apparatus volume within averaged TEM images; IF SPEF2: Immunofluorescence labelling of SPEF2 protein.				

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
P7	0.2	nNO: 1.3 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear and hearing loss NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, some static HSVA post-ALI: static and residual movement TEM: normal Detect: C2b absent IF SPEF2: reduced/absent	Heterozygous variant NM_001270974.2:c.487C > T p. (Arg163Ter)
P8	0.1	nNO: 4.7 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude HSVA post-ALI: reduced coordination and amplitude, semi circling TEM: normal Detect: C2b absent IF SPEF2: reduced	2 heterozygous variants in cis Duplication of exons 22 and 23, deletion of exons 48–50
Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. nNO: nasal nitric oxide; NRDS: neonatal respiratory distress syndrome; HSVA pre-ALI: High speed video microscopy of ciliary beat pattern before ALI culture; HSVA post-ALI: High speed video microscopy of ciliary beat pattern after ALI culture; TEM: Transmission electron microscopy of ciliary ultrastructure; Detect: to visualise central apparatus volume within averaged TEM images; IF SPEF2: Immunofluorescence labelling of SPEF2 protein.				

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
P9*	37	nNO: 20.3 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: nil NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude, rotation HSVA post-ALI: not done TEM: normal Detect: C2b absent IF SPEF2: absent	Heterozygous variant NM_001270974.2:c.3786-1G > T
P10	41	nNO: 23 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: glue ear NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude HSVA post-ALI: reduced coordination and amplitude TEM: normal Detect: C2b present IF SPEF2: reduced/absent	2 Heterozygous variants NM_001270974.2:c.14363_14364del p.(Ser4788TrpfsTer3), NM_001270974.2:c.8973del p.(Glu2992ArgfsTer27) <i>Need confirmation that these are in HYDIN not HYDIN2</i>
P11	8	nNO: nd Cough: daily wet	HSVA pre-ALI: reduced coordination	Heterozygous variant NM_001270974.2:c.12999C > G p.(Tyr4333Ter)

Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. **nNO:** nasal nitric oxide; **NRDS:** neonatal respiratory distress syndrome; **HSVA pre-ALI:** High speed video microscopy of ciliary beat pattern before ALI culture; **HSVA post-ALI:** High speed video microscopy of ciliary beat pattern after ALI culture; **TEM:** Transmission electron microscopy of ciliary ultrastructure; **Detect:** to visualise central apparatus volume within averaged TEM images; **IF SPEF2:** Immunofluorescence labelling of SPEF2 protein.

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
		Situs: normal Rhino-sinus disease: mild Otologic disease: nil NRDS: yes	and amplitude, some static HSVA post-ALI: static and residual movement TEM: normal Detect: C2b absent IF SPEF2: reduced/absent	
P12	18.5	nNO: 6 nL/min Cough: daily wet Situs: normal Rhino-sinus disease: yes Otologic disease: nil NRDS: yes	HSVA pre-ALI: reduced coordination and amplitude HSVA post-ALI: reduced coordination and amplitude TEM: normal Detect: C2b absent IF SPEF2: absent	Heterozygous variant NM_001270974.2:c.14437C > T p. (Arg4813Ter)
P13	3.8	nNO: 13.3 nL/min Cough: daily wet Situs: normal	HSVA pre-ALI: reduced coordination and amplitude HSVA post-ALI: reduced coordination and amplitude,	Heterozygous variants NM_001270974.2:c.1351C > T p. (Arg451Ter)
Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger HYDIN specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. nNO: nasal nitric oxide; NRDS: neonatal respiratory distress syndrome; HSVA pre-ALI: High speed video microscopy of ciliary beat pattern before ALI culture; HSVA post-ALI: High speed video microscopy of ciliary beat pattern after ALI culture; TEM: Transmission electron microscopy of ciliary ultrastructure; Detect: to visualise central apparatus volume within averaged TEM images; IF SPEF2: Immunofluorescence labelling of SPEF2 protein.				

Individual	Age at presentation	Clinical findings	Functional assays	Molecular findings (before RNA-seq)
		Rhino-sinus disease: yes	residual movement TEM: normal	
		Otologic disease: nil	Detect: C2b inconclusive	
		NRDS: yes	IF SPEF2: pre-ALI present, post-ALI present	
Entries marked with an asterisk (*) have been previously published¹⁵ but are presented here as part of a larger <i>HYDIN</i> specific cohort. Probands marked with a dagger (†) are siblings. ‡ Also has an affected sibling with the same variants but did not have RNAseq. nNO: nasal nitric oxide; NRDS: neonatal respiratory distress syndrome; HSVA pre-ALI: High speed video microscopy of ciliary beat pattern before ALI culture; HSVA post-ALI: High speed video microscopy of ciliary beat pattern after ALI culture; TEM: Transmission electron microscopy of ciliary ultrastructure; Detect: to visualise central apparatus volume within averaged TEM images; IF SPEF2: Immunofluorescence labelling of SPEF2 protein.				

RNA-seq genetically solved 11 of 13 *HYDIN* cases. Seven unrelated cases harboured the same pathogenic event. RNA-seq for probands P1, P2, P3, P4, P5, P6, P7 and P8 displayed the same three novel splicing events: exon 18 skipping, two novel acceptor sites and a novel donor site within intron 18 (Fig. 1). Further investigation revealed that these events were caused by a large deletion encompassing part of intron 17, exon 18 and part of intron 18. The breakpoints were established using the soft clipped reads as chr16:71051500–71061418 and validated using Sanger sequencing. All patients had the exact same breakpoints. Long-read RNA sequencing from proband P2 provided orthogonal support for these events. Sashimi plots of long-read alignments demonstrated reads spanning the deletion with splice junctions consistent with exon skipping, cryptic acceptor usage, and pseudoexon inclusion, whereas control samples predominantly exhibited canonical splicing.

Proband P5 was panel negative and as well as the recurrent deletion, we identified a canonical splice variant NM_001270974.2:c.12652-1G > C creating a novel acceptor site within intron 74 (Fig. 2). In two other probands (P9 and P12) with only one heterozygous variant previously detected, we identified the second event through aberrant splice signals. P9 showed exon 27 skipping, later found by whole genome sequencing to be linked to a deletion of exon 27 [previously published]¹⁵. P12 showed upregulation of exon 21 skipping caused by a canonical splice variant NM_001270974.2:c.3043-1G > A. While the read coverage of the splice variant was low, the variant was Sanger validated.

Lastly, proband P10 required confirmation that both variants previously identified through standard genomic testing were in fact in *HYDIN* and not *HYDIN2*. Since *HYDIN2* transcripts are not present in airway cells and both variants were present in the RNA-seq, we could confirm these were indeed in *HYDIN*.

Discussion

RNA-seq enabled us to identify and confirm pathogenic *HYDIN* splice-altering variants, resolving the molecular diagnosis in this 'TEM normal' cohort with high clinical suspicion of *HYDIN*-related PCD. Our cohort was enriched with probands harbouring PCD gene-panel identified heterozygous pathogenic or likely pathogenic variants in *HYDIN*, but lacking evidence of another biallelic event for a definitive molecular diagnosis. This enrichment reflects the nature of cases currently referred through routine diagnostic pathways, rather than a retrospectively assembled *HYDIN* specific cohort. The data show that RNA-seq is particularly valuable for identifying the recurrent *HYDIN* deletion (c.2376+752_2529 + 9004del), which we now describe as a *HYDIN* 'hotspot', accounting for over half of the cases in our cohort. Furthermore, RNAseq identified both *HYDIN* variants in a proband with no initial candidates suggesting there is utility in including patients with a strong clinical suspicion even when genomic DNA testing is negative.

A number of patients demonstrated presence or reduced presence of anti-SPEF2 localisation to cilia, in particular patients P6 (original biopsy) and P13 with immunofluorescence labelling intensity similar to healthy donor control samples. Importantly, P6 carries a missense variant in addition to the *HYDIN* hotspot deletion, while P13 currently remains without a molecular diagnosis. Reinforcing that despite its utility in aiding a highly likely diagnosis of PCD, 'false negative' normal presence of SPEF2 localisation can occur. Normal axonemal localisation of SPEF2 has been shown to be dependent on the function of *HYDIN* and its protein, meaning immunofluorescent labelling of SPEF2 can occur with hypomorphic *HYDIN* variants⁶. This observation may extend more widely to IF assays across other PCD genes, where missense or hypomorphic variants may preserve protein localisation despite pathogenic effects on ciliary function. Similarly, normal nNO has also been reported in a subset of PCD patients harbouring *HYDIN* variants⁹, highlighting the complexities of diagnosis PCD caused by *HYDIN* defects.

The *HYDIN* 'hotspot' deletion has been previously reported in an individual from the Genomics England cohort²⁷ and in two affected sibling reported by Fleming *et al.* found in *trans* with a likely pathogenic variant identified through targeted long read DNA sequencing⁹. Notably, presence of SPEF2 immunofluorescence was reported in one of the siblings, in line with our results, showing reduced or present staining in patients harbouring this hotspot. This deletion is missed by standard genomic DNA-based approaches and since skipping of exon 18 is predicted to result in an in-frame deletion. The variant would be classified as a VUS based on the genomics results alone. However, RNA-seq demonstrates that the predominant aberrant transcripts are out of frame, providing functional evidence that supports reclassification of the variant to likely pathogenic. While this variant is not novel, its high frequency has not been previously recognised. All carriers of this hotspot have the same breakpoint suggesting that this event arose from a single ancestral event and is identical by descent. Our findings suggest that *HYDIN*-related PCD may be more common than current estimates indicate and could represent one of the more frequent genetic causes of PCD within the UK. The deletion was likely missed by standard testing, due to deep intronic breakpoints and high homology to *HYDIN2*. However, we now have a DNA-based targeted assay that can be incorporated into routine diagnostic testing.

While previous work by Hijikata et. al⁸ used library enrichment to obtain sufficient depth of *HYDIN* transcripts, this work demonstrates that specific enrichment is not needed to successfully capture this gene in the relevant tissue where it is abundantly expressed. Furthermore, previous studies have shown that *HYDIN2* is not expressed in airway cells⁷, therefore all the transcripts identified should belong to *HYDIN*. This eliminates the ambiguity of whether a variant is present in *HYDIN* or *HYDIN2* and the need for long read sequencing to determine provenance. However, it might be that RT-PCR in airway tissue might be enough to establish provenance as shown by Burgoyne and colleagues²⁸.

There is no perfect test to diagnose PCD and RNA-seq is no exception. Clinical information and ciliary function testing provide essential context that helps prioritise relevant variants and interpret RNA-seq findings more accurately. Using RNA-seq alone to identify non-coding variants, or aberrant splicing has its own limits as we will miss variants that have been spliced out unless there are pre-spliced reads present (sometimes observed). Additionally, splicing events that created truncated transcripts might be missed if the affected transcript has undergone nonsense mediated decay. Structural variants are hard to infer from RNA-seq data²⁹ and while we were able to observe the breakpoints in the recurrent hotspot deletion, this is not usually the case. Finally, RNA variant calling has a much lower sensitivity than DNA based approaches, which leads to labour intensive and costly manual variant curation¹⁵.

RNA-seq is a key tool to uplift genetic diagnosis in rare genetic disorders and should be implemented in the diagnosis of PCD particularly those suspected to have *HYDIN* defects. Its use on routinely collected samples removes the need for additional patient procedures, which could encourage more patients to opt into this type of study which is currently research based. Importantly, RNA-seq enables detection of noncoding and splice- altering variants that are frequently missed by DNA-based testing and provides essential functional evidence to support confident variant classification.

Declarations

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability

The code generated during this study can be found on GitHub: <https://github.com/TheRealLBaker/pcd-transcriptomics>

Acknowledgements

The authors thank all patients and families taking part in this research and all clinicians involved. The authors acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work.

Author contributions

Conceptualization: C.L.J., C.J.O.; Data curation: C.J.O, J.C., L.B., D. H.; Formal analysis: C.J.O., J.C., D.J.B.; Funding acquisition: C.L.J., D.B., J.L., J.H.; Investigation: C.J.O., L.B., J.C., J.T., D.J.B., G.F., S.K.; Methodology: C.J.O., L.B., J.C.; Project Administration: J.C., C.J.O.; Resources: C.L.J., N.S.T., W.T.W.; Software: L.B.; Supervision: C.L.J., N.S.T., W.T.W.; Validation: N.S.T., D.J.B.; Visualisation: C.J.O., J.C., J.T.; Writing-original draft: C.J.O., J.C., C.L.J.; Writing-review & editing: All authors

Funding

Funded National Research Project Grants | GOSH Charity “Uplifting genetic detection of rare lung diseases; a novel diagnostic pipeline utilising RNA-seq analysis” awarded in 2024 to C.L.J and supplying funding for consumables, RNA-seq services and C.J.O post.

AAIR Research Grant “Mapping RNA isoform diversity in nasal epithelial cells to improve diagnostics and understanding of disease severity in Primary Ciliary Dyskinesia Patients” awarded in 2025 to C.J.O and J.C for consumables.

The Southampton PCD Service is commissioned and funded by NHS England.

Ethical Approval

Patients were ethically consented to the HRA approved “A study of the pathophysiology and clinical manifestations of Primary Ciliary Dyskinesia” LREC 06/Q1702/109: CHI 0395, sponsored by the University Hospital Southampton NHS Foundation Trust.

Competing Interests

The authors declare that they have no competing interests.

References

1. Hannah WB *et al.* The global prevalence and ethnic heterogeneity of primary ciliary dyskinesia gene variants: a genetic database analysis. *The Lancet Respiratory Medicine* 2022; 10: 459–468.
2. Goutaki M *et al.* Clinical manifestations in primary ciliary dyskinesia: systematic review and meta-analysis. *Eur Respir J* 2016; 48: 1081–1095.
3. Shoemark A *et al.* European Respiratory Society and American Thoracic Society guidelines for the diagnosis of primary ciliary dyskinesia. *European Respiratory Journal* 2025; 66. doi:10.1183/13993003.00745-2025.
4. Shoemark A *et al.* International consensus guideline for reporting transmission electron microscopy results in the diagnosis of primary ciliary dyskinesia (BEAT PCD TEM Criteria). *European Respiratory Journal* 2020; 55. doi:10.1183/13993003.00725-2019.

5. Marshall CR et al. Whole-Exome Sequencing and Targeted Copy Number Analysis in Primary Ciliary Dyskinesia. *G3 (Bethesda)* 2015; 5: 1775–1781.
6. Cindrić S et al. SPEF2- and HYDIN-Mutant Cilia Lack the Central Pair-associated Protein SPEF2, Aiding Primary Ciliary Dyskinesia Diagnostics. *Am J Respir Cell Mol Biol* 2020; 62: 382–396.
7. Dougherty ML et al. The birth of a human-specific neural gene by incomplete duplication and gene fusion. *Genome Biol* 2017; 18: 49.
8. Hijikata M, Morimoto K, Ito M, Wakabayashi K, Miyabayashi A, Keicho N. Robust detection of pathogenic HYDIN variants that cause primary ciliary dyskinesia using RNA-seq of nasal mucosa. *Journal of Medical Genetics* 2025; 62: 180–184.
9. Fleming A et al. Combined approaches, including long-read sequencing, address the diagnostic challenge of HYDIN in primary ciliary dyskinesia. *Eur J Hum Genet* 2024; 32: 1074–1085.
10. Yépez VA et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Medicine* 2022; 14: 38.
11. Murdock DR et al. Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of conventional genomic testing. *J Clin Invest* 2021; 131: e141500.
12. Jaramillo Oquendo C et al. Identification of diagnostic candidates in Mendelian disorders using an RNA sequencing-centric approach. *Genome Medicine* 2024; 16: 110.
13. Gonorazky HD et al. Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. *The American Journal of Human Genetics* 2019; 104: 466–483.
14. Cummings BB et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med* 2017; 9: eaal5209.
15. Legebeke J et al. Uplift of genetic diagnosis of rare respiratory disease using airway epithelium transcriptome analysis. *Hum Mol Genet* 2025; 34: 148–160.
16. Beydon N et al. Nasal nitric oxide measurement in children for the diagnosis of primary ciliary dyskinesia: European Respiratory Society technical standard. *Eur Respir J* 2023; 61: 2202031.
17. Coles JL et al. A Revised Protocol for Culture of Airway Epithelial Cells as a Diagnostic Tool for Primary Ciliary Dyskinesia. *Journal of Clinical Medicine* 2020; 9. doi:10.3390/jcm9113753.
18. Shoemark A et al. PCD Detect: enhancing ciliary features through image averaging and classification. *Am J Physiol Lung Cell Mol Physiol* 2020; 319: L1048–L1060.
19. Durkie M et al. ACGS best practice guidelines for variant classification in rare disease 2024. *Association for clinical genomic science* 2024; 56.
20. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29: 15–21.
21. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26: 841–842.
22. Wang Y, Xie Z, Kutschera E, Adams JI, Kadash-Edmondson KE, Xing Y. rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. *Nat Protoc* 2024; 19: 1083–1104.

23. Robinson JT *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011; 29: 24–26.
24. Garrido-Martín D, Palumbo E, Guigó R, Breschi A. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS Comput Biol* 2018; 14: e1006360.
25. PacificBiosciences/pbmm2. 2026.<https://github.com/PacificBiosciences/pbmm2> (accessed 21 Apr2026).
26. Tang AD, Felton C, Hrabeta-Robinson E, Volden R, Vollmers C, Brooks AN. Detecting haplotype-specific transcript variation in long reads with FLAIR2. *Genome Biol* 2024; 25: 173.
27. Wheway G *et al.* Whole genome sequencing in the diagnosis of primary ciliary dyskinesia. *BMC Med Genomics* 2021; 14: 234.
28. Burgoyne T *et al.* HYDIN variants cause primary ciliary dyskinesia in the Finnish population. *Pediatric Pulmonology* 2024; 59: 3601–3609.
29. Xiao B *et al.* Combining optical genome mapping and RNA-seq for structural variants detection and interpretation in unsolved neurodevelopmental disorders. *Genome Med* 2024; 16: 113.

Figures

Central pair complex

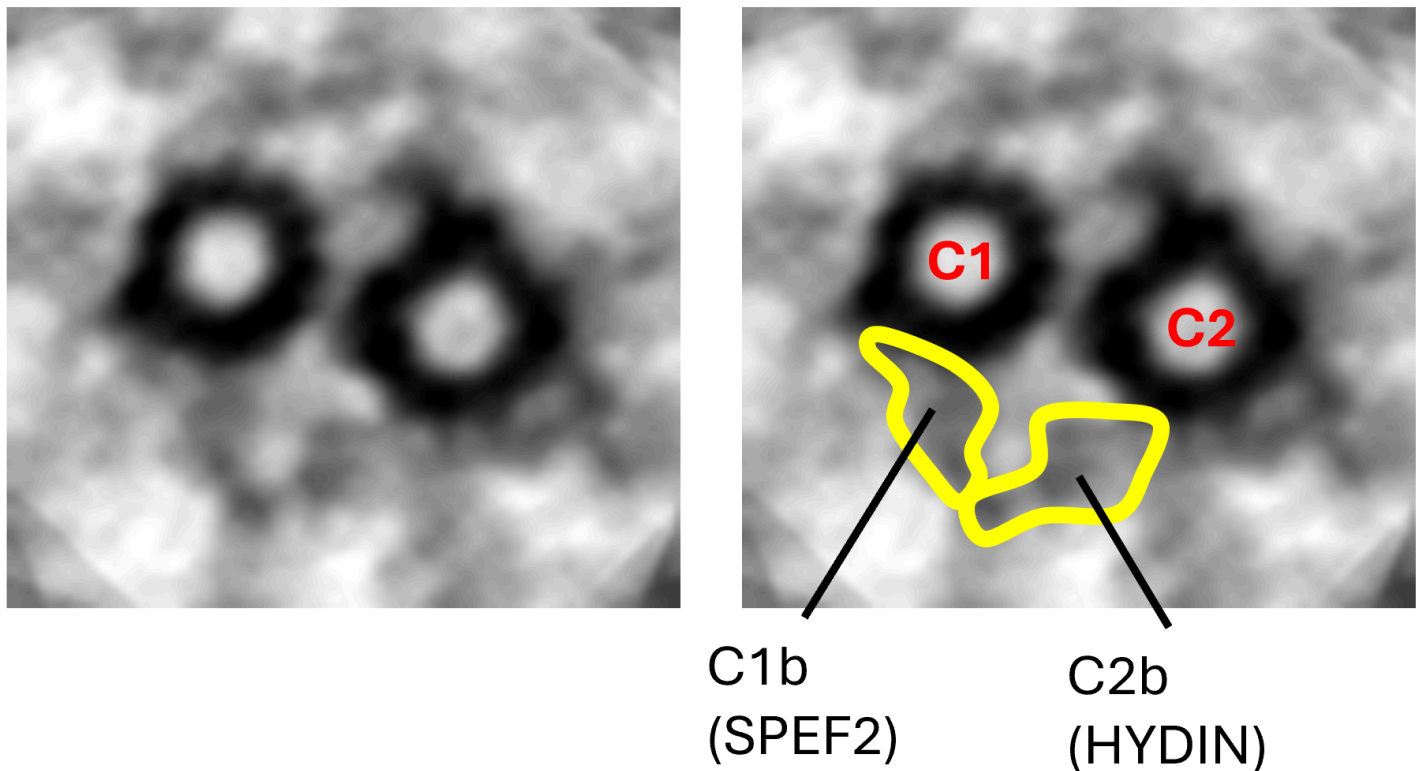


Figure 1

TEM Cross-section of a normal central pair complex. C1 and C2 central microtubules are indicated in red. Central pair projections are outlined in yellow, with arrows marking the C1b projection (SPEF2) and the C2b projection (HYDIN).

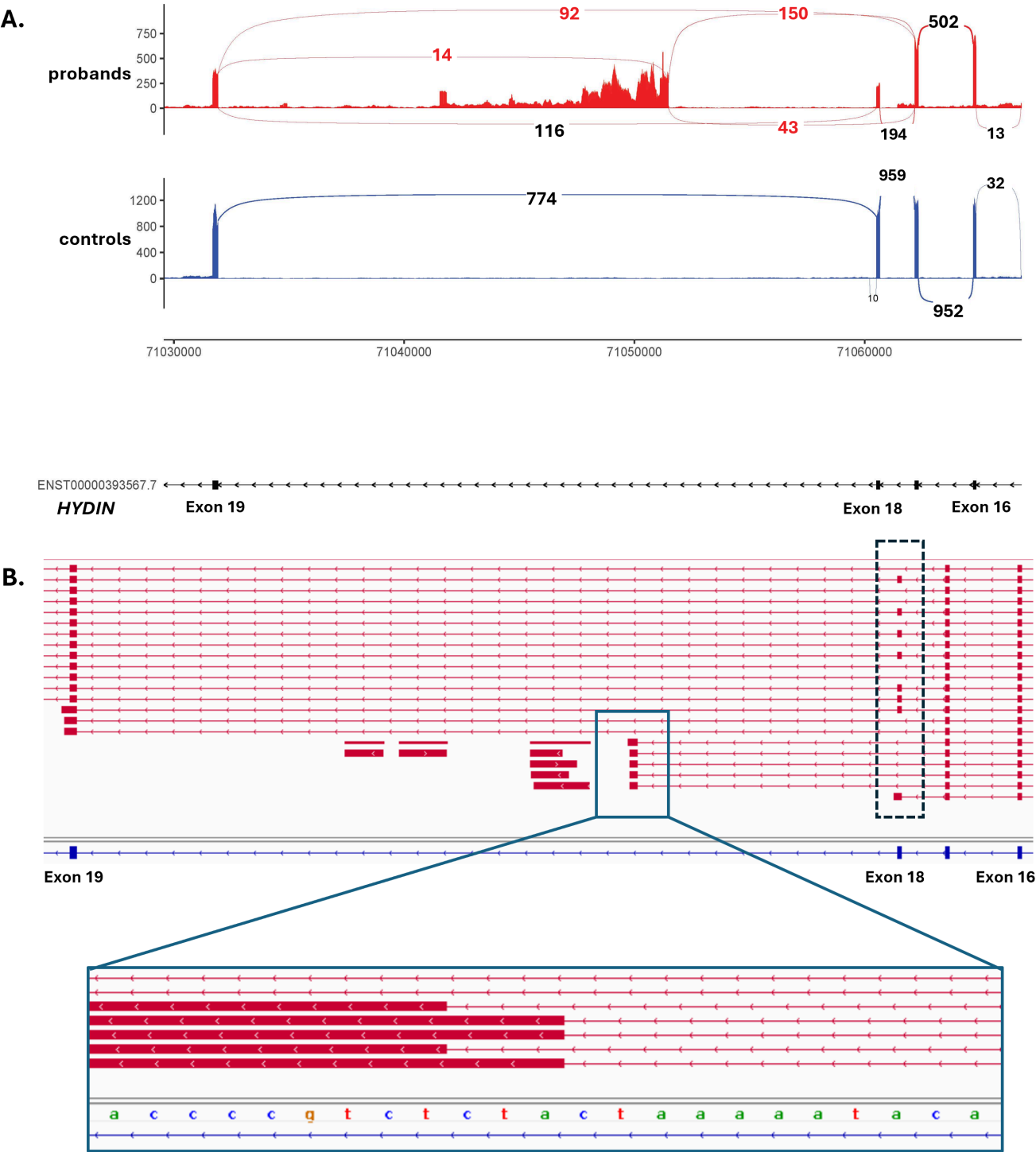


Figure 2

Effects of the recurrent *HYDIN* deletion (chr16:71051501-71061418) on mRNA. **A.** The sashimi plot shows the observed splicing patterns between probands (red n=8) and controls (blue n=9). The arcs represent the junctions present within the data and the number the median number of reads across the junctions. The probands all show multiple events absent in controls including exon 18 skipping, two novel acceptor sites and a novel donor site within intron 18. Numbers in red indicate novel splice junctions. **B.** FLAIR-collapsed full-length transcript isoforms derived from long-read RNA sequencing (red). Each horizontal line represents a distinct transcript isoform, with boxes indicating exons and connecting lines indicating introns. The dashed box highlights the exon skipping event, while the zoomed inset details the two novel acceptor sites within intron 18. The gene model and exon numbering (blue) are shown below for reference.

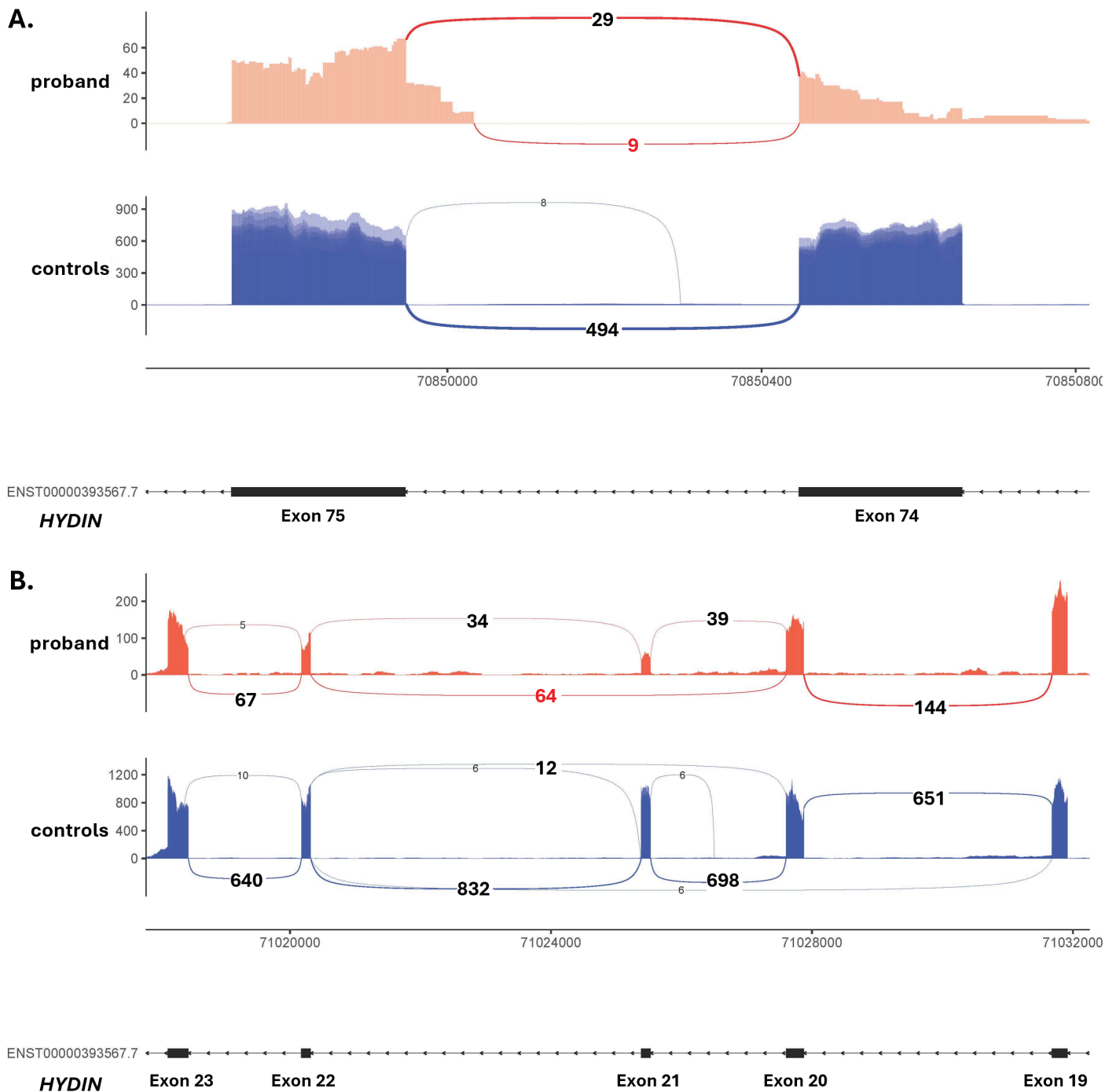


Figure 3

Aberrant splicing events detected by RNA-seq. **A.** Activation of novel acceptor site in patient P5 caused by canonical splice variant NM_001270974.2:c.12652-1G>C. **B.** Upregulation of exon 21 in patient P12 caused by a canonical splice variant NM_001270974.2:c.3043-1G>A. Numbers in red indicate novel splice junctions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementalfigure1.tif](#)