RESEARCH ARTICLE

Human Mutation | HGVS HUMAN GENOME VARIATION SOCIETY | WILEY

High-yield identification of pathogenic NF1 variants by skin fibroblast transcriptome screening after apparently normal diagnostic DNA testing

Hannie C. W. Douben[1] | Mark Nellist[1] | Leontine van Unen[1] | Peter Elfferich[1] |
Esmee Kasteleijn[1] | Marianne Hoogeveen-Westerveld[1] | Jesse Louwen[1] |
Monique van Veghel-Plandsoen[1] | Walter de Valk[1] | Jasper J. Saris[1] |
Femke Hendriks[1] | Esther Korpershoek[1,2] | Lies H. Hoefsloot[1] |
Margreethe van Vliet[1,3] | Yolande van Bever[1] | Ingrid van de Laar[1] |
Emmelien Aten[4] | Augusta M. A. Lachmeijer[5] | Walter Taal[3,6] |
Lisa van den Bersselaar[1] | Juliette Schuurmans[5] | Rianne Oostenbrink[3,7] |
Rick van Minkelen[1,3] | Yvette van Ierland[1,3] | Tjakko J. van Ham[1]

[1]Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[2]Department of Pathology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[3]ENCORE Expertise Center for Neurodevelopmental Disorders, Erasmus MC, Rotterdam, The Netherlands

[4]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands

[5]Department of Genetics, Division Laboratories, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

[6]Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands

[7]Department of General Pediatrics, Erasmus MC—Sophia Children's Hospital, Rotterdam, The Netherlands

Correspondence
Tjakko J. van Ham, Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands.
Email: t.vanham@erasmusmc.nl

Abstract

Neurofibromatosis type 1 (NF1) is caused by inactivating mutations in NF1. Due to the size, complexity, and high mutation rate at the NF1 locus, the identification of causative variants can be challenging. To obtain a molecular diagnosis in 15 individuals meeting diagnostic criteria for NF1, we performed transcriptome analysis (RNA-seq) on RNA obtained from cultured skin fibroblasts. In each case, routine molecular DNA diagnostics had failed to identify a disease-causing variant in NF1. A pathogenic variant or abnormal mRNA splicing was identified in 13 cases: 6 deep intronic variants and 2 transposon insertions causing noncanonical splicing, 3 postzygotic changes, 1 branch point mutation and, in 1 case, abnormal splicing for which the responsible DNA change remains to be identified. These findings helped resolve the molecular findings for an additional 17 individuals in multiple families with NF1, demonstrating the utility of skin-fibroblast-based transcriptome analysis for molecular diagnostics. RNA-seq improves mutation detection in NF1 and

provides a powerful complementary approach to DNA-based methods. Importantly, our approach is applicable to other genetic disorders, particularly those caused by a wide variety of variants in a limited number of genes and specifically for individuals in whom routine molecular DNA diagnostics did not identify the causative variant.

1 | INTRODUCTION

Neurofibromatosis type 1 (NF1) is an autosomal dominant tumor predisposition syndrome caused by inactivating mutations in the *NF1* tumor suppressor gene that encodes neurofibromin, a RAS GTPase activating protein (GAP) (Ballester et al., 1990; Trovó-Marqui & Tajara, 2006). NF1 is characterized by the occurrence of multiple café-au-lait (CAL) spots, skinfold freckling, neurofibromas, eye and skeletal abnormalities, and developmental and/or behavioral problems (Jett & Friedman, 2010). NF1-associated tumors can become malignant, and, therefore, establishing a diagnosis of NF1 is important for guiding treatment and monitoring. In the majority of individuals with NF1, a clinical diagnosis can be made after the age of 6 years, according to consensus-based diagnostic criteria (Legius et al., 2021). However, in young or mildly affected individuals with no affected first-degree relative, establishing a clinical diagnosis is more difficult and molecular testing is extremely useful for confirmation of a suspected diagnosis. In addition, the identification of a (likely) pathogenic *NF1* variant facilitates family counseling and prenatal diagnostic testing.

NF1 is a large gene, consisting of 61 exons that span >350 kb of genomic DNA on chromosome 17. The size and complexity of the *NF1* locus, as well as the occurrence of mosaicism (Colman et al., 1996; Lazaro et al., 1994), make molecular diagnostics challenging. Indeed, in ~10%–20% of individuals meeting diagnostic criteria for NF1, DNA-based screens fail to identify a pathogenic variant (van Minkelen et al., 2014; Pasmant et al., 2015). These individuals are often referred to as NF1 "no mutation identified" (NMI). In some laboratories, particularly those with a focus on NF1, routine DNA screening by polymerase chain reaction (PCR)/Sanger sequencing and multiplex ligation probe amplification (MLPA) has been complemented or replaced entirely by RNA- and next-generation sequencing (NGS)- based techniques that increase the diagnostic yield to >95% by identifying variants that affect premRNA splicing and/or are present at a low frequency due to mosaicism (Ars et al., 2000; Evans et al., 2016; Koster et al., 2021; Martinez et al., 1996; Messiaen et al., 2000; Serra et al., 2001; Wimmer et al., 2011, 2020; Zatkova et al., 2004). Unfortunately, such a specialized, focused approach is not available in all diagnostic laboratories. Furthermore, there is a need for simple, complementary approaches that can be applied to increase the molecular diagnostic yield, not just in NF1 but also in other genetic conditions.

Transcriptome analysis of RNA from skin fibroblasts by high throughput sequencing of cDNA libraries—RNAseq—, has been shown to be effective for the detection of pathogenic splice variants missed by routine DNA sequencing in a range of disorders (Cummings et al., 2017; Dekker et al., 2022; Kremer et al., 2017) and we reasoned that RNAseq of skin fibroblast RNA might improve molecular diagnostics for NF1 in our laboratory. Here we present our initial findings in a cohort of 15 NF1 NMI individuals. We demonstrate the utility of RNAseq for the identification, classification, and characterization of pathogenic *NF1* variants in 13 individuals meeting diagnostic criteria for NF1 (87% of the cohort).

2 | MATERIALS AND METHODS

2.1 | Patient assessment and selection of variants for testing

Individuals who were *NF1* NMI after routine DNA-based molecular screening (van Minkelen et al., 2014) and met diagnostic criteria for NF1, and for whom a molecular diagnosis was considered useful or important, were included in the study. Clinical and genetic information was extracted from the Erasmus MC Department of Clinical Genetics NF1 patient database. Nomenclature for all reported variants was according to the current Human Genome Variation Society guidelines and NM_000267.3(NF1) is used throughout this manuscript as the reference transcript. Variants were classified according to American College of Medical Genetics guidelines (Richards et al., 2015) using splice prediction software (Alamut Visual v.2.15; Sophia Genetics) and the available clinical, genetic, and functional data. Clinical characteristics are summarized in Supporting Information: Table S1. Routine DNA-based molecular screening of *NF1* and, in some cases, *SPRED1* by Sanger sequencing and MLPA did not identify the causative variant in any of the individuals in the cohort. All individuals or their legal representatives provided verbal and/or written consent for diagnostic testing.

2.2 | Transcriptome analysis

For isolation of total RNA from skin-derived fibroblast cultures, cells were cultured at 37°C in Ham's F10 (15% [v/v] fetal calf serum), with and without cycloheximide (250 μg/ml; overnight), in 75 cm² flasks. RNA was isolated using the RNeasy mini kit (Qiagen). Quality was assessed using a Bioanalyzer (Agilent), and cDNA synthesis was performed using NEBNext PolyA and NEBNext RNA Ultra II Directional kit (NEB). RNA-seq was carried out by GenomeScan on an Illumina NovaSeq. 6000 (150 bp paired-end reads including unique molecular identifier [UMI]-adaptors, >40 million total unique reads per sample). The FASTQ datasets were processed at the Department of Clinical Genetics (Erasmus MC). Reads were merged per sample and mapped using HiSat2 (v2.1.0) (Kim et al., 2019). HiSat2 was used to align sequence reads to sequence alignment map (SAM) format. Samtools (v1.9) was used to convert SAM to binary alignment map format and to mark duplicate reads, followed by Bedtools (v2.29.2) to count the reads per coding sequence (CDS). Only uniquely mapped reads with MAPQ scores < 1 were removed from the counts (Wang et al., 2016). The reference genome GRCh37/hg19 was used. Data were visualized using the Integrated Genome Viewer (IGV, 2.11.17) (Robinson et al., 2011).

2.3 | RT-PCR; Sanger sequencing; allele-specific (AS)-PCR; functional assessment

All variants were verified in the corresponding genomic DNA by PCR followed by Sanger sequencing. RNA-seq data were confirmed by standard RT-PCR followed by agarose gel electrophoresis and Sanger sequence analysis. NF1 wild-type and variant-specific primers for AS-PCR were designed using Primer 3 (https://primer3.ut.ee/). Minigene exon trap constructs were generated in the pSPL3 vector by Gibson assembly (Church et al., 1994; Gibson et al., 2009). The NF1 variant constructs used for this study were derived by site-directed mutagenesis using the QuikChange II XL kit (Agilent). All constructs were verified by Sanger sequencing. Minigene constructs were transfected into HEK 293T cells using polyethyleneimine and, after 24–48 h, cDNA was prepared using the iScript cDNA synthesis kit (BioRad), as described previously (Dufner-Almeida et al., 2020). RT-PCR products were analyzed by agarose gel electrophoresis and Sanger sequencing. Primer sequences for AS-PCR, Gibson cloning, RT-PCR, and Sanger sequencing are available on request.

3 | RESULTS

We performed RNA-seq on RNA isolated from fibroblasts obtained from skin biopsies of 15 NF1 NMI individuals. In each case, routine DNA-based molecular screening had failed to identify a pathogenic variant in NF1. Eleven individuals fulfilled the diagnostic criteria for NF1 (Legius et al., 2021), and mosaicism was suspected in five individuals. The clinical findings of the patients are summarized

in Supporting Information: Table 1. Fibroblasts from each individual were either treated with cycloheximide to inhibit nonsense-mediated mRNA decay (NMD), or left untreated, before RNA extraction, and showed robust levels of NF1 expression (Supporting Information: Figure S1). Total cellular RNA was subjected to polyA-based cDNA library synthesis and sequencing (see Section 2 for details). Reads were mapped to the human genome (GRCh37/hg19) using HiSat2 and visualized in the Integrated Genome Viewer (IGV 2.4.15; Broad Institute). The results are summarized in Table 1. Variant nomenclature is according to reference transcript NM_000267.3(NF1), unless specified otherwise. The standard molecular diagnostic screening had identified at least one benign heterozygous single-nucleotide polymorphism (SNP) in an exon of NF1 in all individuals in our cohort. This enabled us to quickly establish that 9/15 individuals (60%) showed evidence for skewed or mono-allelic NF1 expression. For these individuals, we focused our search on possible germ-line mutations. In six cases, NF1 expression was bi-allelic, suggesting that these individuals were more likely to be mosaic for a pathogenic NF1 variant.

In six individuals (1–6) we identified noncanonical pseudoexons mapping to NF1 introns 13, 8 (2×), 14, and 39 (2×) (Figure 1a–c, Supporting Information: Figure S2). In all cases, the abnormal transcripts were likely subject to NMD as they were most clearly visible in the data set corresponding to the cycloheximide-treated fibroblasts.

Subsequent Sanger DNA-sequencing of the intronic segments surrounding the pseudoexons led to the identification of the heterozygous deep-intronic variants c.1527 + 1187C>G, c.888 + 789A>G (2×), c.1642 − 449A>G, c.5749 + 267A>G and c.5749 + 332A>G that were all absent from gnomAD (v2.1.1) (Table 1). In silico splice site analysis predicted that the variants either created or strengthened a noncanonical donor or acceptor site that, in combination with a nearby noncanonical acceptor or donor site resulted in the incorporation of the corresponding pseudoexon into the NF1 transcript (Figure 1a–c, Supporting Information: Figure S2). The c.888 + 789A>G and c.5749 + 332A>G variants were reported previously as likely pathogenic (Class 4) in individuals with NF1 (Evans et al., 2016; Pros et al., 2008; Sabbagh et al., 2013).

In individual 7, we identified an unexpected stretch of transcribed sequence immediately following exon 15 (~30% total reads) in both untreated and cycloheximide-treated RNA fractions, equivalent to reference transcript NM_001128147.2(NF1) that encodes neurofibromin isoform 3 (Figure 2a). This truncated isoform lacks the RAS GAP, SPRED1 interaction, and SEC14 domains that are critical for neurofibromin function (Suzuki et al., 1992). In addition to the reads mapping to NF1 intron 15, the read depths of the exons distal to exon 15 were reduced relative to the proximal exons and to control samples. Furthermore, based on the variant allele fractions (VAF) for SNPs in exons 7 and 18, NF1 expression appeared skewed (Table 1, data not shown). We suspected that a novel DNA variant identified adjacent to exon 16 in this individual, c.1722-26T>G, might be responsible for the apparent intron retention. In silico analysis suggested that this variant might disrupt the canonical exon 16

TABLE 1 Abnormal NF1 expression and mRNA splicing, and pathogenic NF1 variants identified in NF1 NMI individuals

Summarized findings

Patient	NF1 variant NC_000017.10(NM_000267.3) (position)	Effect on transcript	Type mutation/effect	Validation	Detected previously/freq.
1	c.1527 + 1187C>G (i13)	r.1527_1528ins[1527 + 1104_1527 + 1182]	Creates 5′ donor; Pseudo-exon, skewed expr.	PCR	No
2	c.888 + 789A>G (i8)	r.888_889ins[888 + 656_888 + 784], r.888_889ins [888 + 710_888 + 784]	Strengthens noncanonical 5′ donor; Pseudo-exon, skewed expr.	PCR	ClinVar: 664551 (Evans et al., 2016)
3	c.888 + 789A>G (i8)	r.888ins888 + 656_888 + 784/ r.888ins888 + 710_888 + 784	Strengthens noncanonical 5′ donor; Pseudo-exon, skewed expr.	PCR	ClinVar: 664551 (Evans et al., 2016)
4	c.1642 − 449A>G (i14)	r.1641ins1642-448_1642-405, p.(Ala548Glufs*11)	Creates 3′ acceptor; Pseudo-exon	PCR	ClinVar: 216394
5	c.5749 + 332A>G (i39)	r.5749ins5749+156_5949+331, p.(Ser1917Lysfs*3)	Creates 5′ donor; Pseudo-exon	PCR	ClinVar: 216065
6	c.5749 + 267A>G (i39)	r.5749ins5749+156_5949+267, p.(Ser1917Lysfs*3)	Creates 5′ donor; Pseudo-exon	PCR	No
7	c.1722 − 26T>G (i15)	r.1722_1845del, p.(Ser575Argfs*15)	Branch point/activation isoform 3, skewed expr.	exon trapping	No
8	c.3312_3313ins[AluYa5;3299_3312dup] (e25)	skipping e23-25: r.2991_3314del, r.3275_3314del, r.3211_3314del	Transposon ins. (AluYa5), skewed expr.	PCR	No
9	c.6852_6853ins[T[?];AluY;6839_6852dup] (e46)	r.6757_6858del, p.(Ala2253_Lys2286del)	Transposon ins. (AluY), skewed expr.	PCR	No
10	c.4265C>T, p.(Ser1422Leu) (e32)	r.4265c>u	Missense, bi-allelic expr.	AS-PCR	Class 5 ClinVar: 527535
11	c.3916C>T, p.(Arg1306*) (e29)	r.3916c>u	Stop gain, bi-allelic expr.	AS-PCR	Class 5 ClinVar: 404592
12	c.7043G>A, p.(Trp2348*) (e48)	Premature stop	Premature stop	AS-PCR	Class 5 ClinVar: 996432
13	Skipping exon 25, NMI (e25)	confirmed low fraction abnormal splicing	Possibly postzygotic, bi-allelic expr.	n.a.	n.a.
14	NMI	n.a.	Possibly postzygotic, bi-allelic expr.	n.a.	n.a.
15	NMI	n.a.	Possibly postzygotic, bi-allelic expr.	n.a.	n.a.

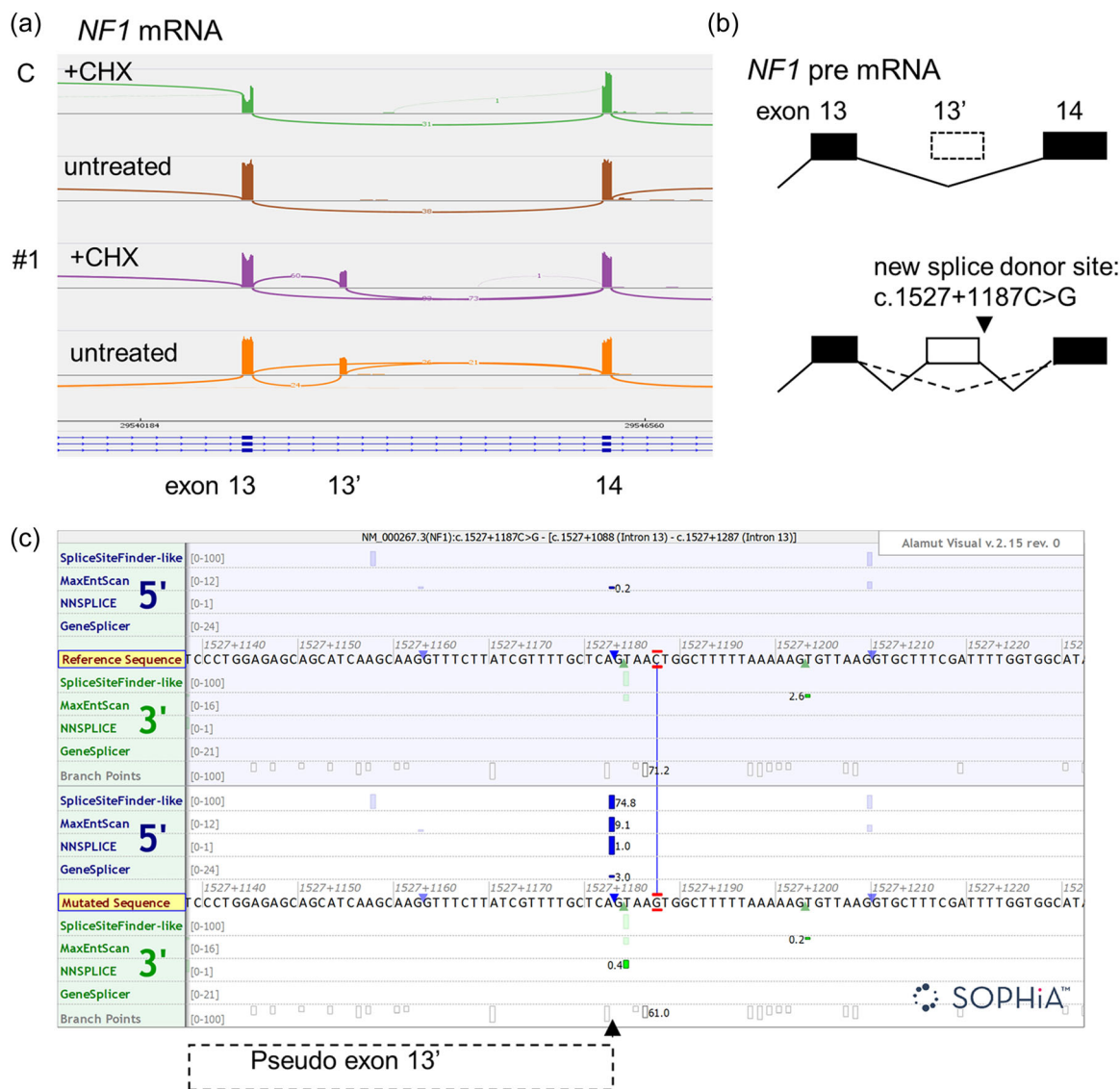Abbreviations: AS-PCR, allele-specific PCR; e/i, exon/intron.

FIGURE 1 RNA-seq analysis of fibroblast mRNA identifies deep intronic. (a) Sashimi plot of control and NF1 individual 1 showing a pseudoexon in intron 13. (b) Schematic showing the position of the variant likely to be causative for the inclusion of the pseudoexon. (c) Alamut screenshot showing splice site predictions for variant c.1527 + 1187C>G. The predicted noncanonical 5′ donor site (blue blocks) and 3′ extent of the detected pseudoexon are indicated.

branch point (Figure 2c). We tested the effect of the c.1722-26T>G variant on pre-mRNA splicing in a mini-gene exon trap assay, which revealed that the variant induced skipping of exon 16, resulting in NM_000267.3(NF1):r.1722_1845del, p.(Ser575Argfs*15) (Figure 2c). We did not detect skipping of exon 16 in the RNA-seq data for individual 7 (Figure 2c) and hypothesize that disruption of the canonical exon 16 branch point favors the expression of the truncated NM_001128147.2(NF1) transcript, rather than skipping of exon 16. To determine whether other variants in this region might have a similar effect on NF1 pre-mRNA splicing, we searched our NF1 patient database and identified 3 variants: c.1722-24A>G, c.1722-14_1722-5del10, and c.1722-16_1727del22. In each case, we confirmed that the variant prevented trapping of exon 16 in the minigene assay and therefore could cause a switch toward the

expression of the truncated NM_001128147.2(NF1) transcript in vivo (Figure 2c, Table 1). Remarkably, the c.1722-24A>G variant was absent from gnomAD, but was identified in 14 NF1 NMI individuals from four unrelated families in our NF1 cohort. All four variants (c.1722-26T>G, c.1722-24A>G, c.1722-14_1722-5del10, and c.1722-16_1727del22) are likely pathogenic, after RNA analysis, based on ACMG criteria: PS3, PM2, PM4, PP1, PP4.

In individuals 8 and 9, we identified specific, noncanonical exon skipping events as well as multiple reads mapping to an exon in NF1 and to repetitive sequences outside the NF1 locus (Figure 3a,b, Table 1). Detailed analysis of these reads indicated that they represented AluY repeat element insertions in exons 25 and 45, respectively. Exon skipping was confirmed by RT-PCR followed by Sanger sequencing, and the insertions were confirmed by PCR
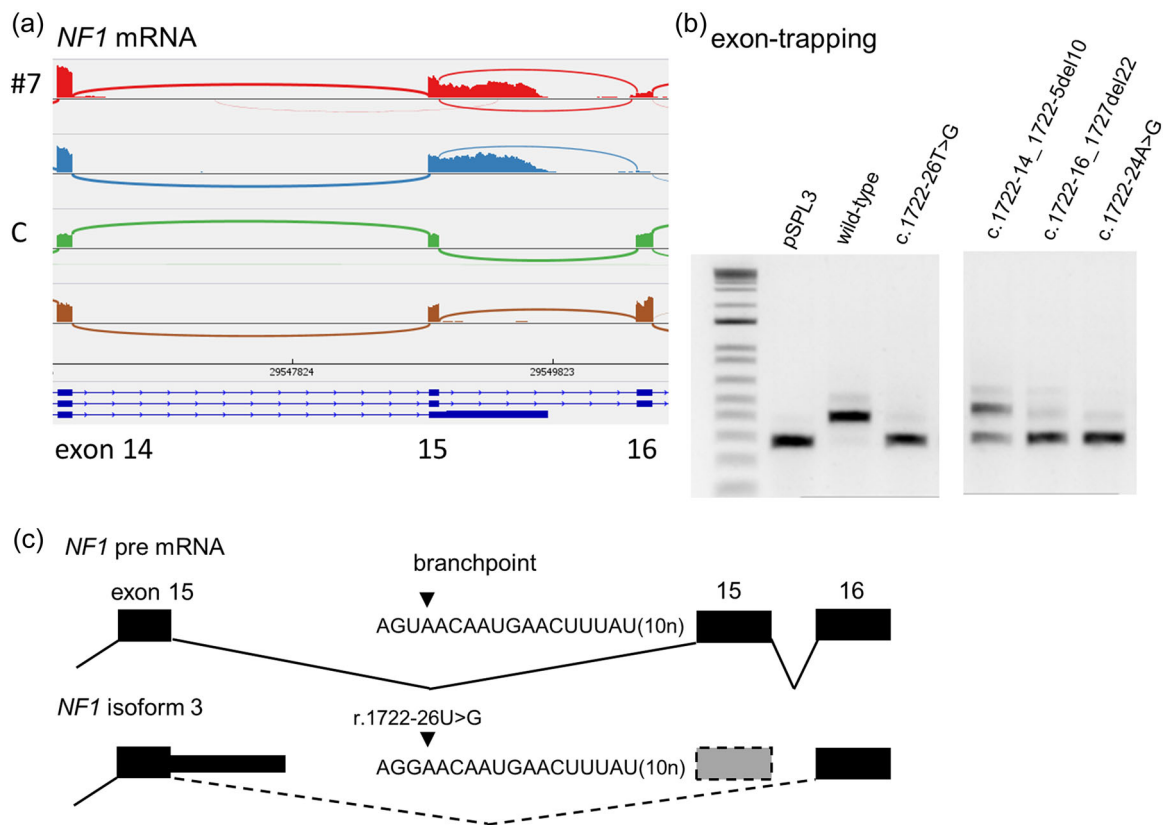
FIGURE 2 RNA-seq analysis of fibroblast mRNA identifies *NF1* variants favoring expression of isoform 3. (a) Sashimi plot of control and NF1 individual 7. (b) Agarose gel analysis of exon trapping experiments showing the effects of variants identified in *NF1* intron 15; (pSPL3). (c) Schematic indicating the position of the c.1722-26T>G variant with respect to the branch point proximal to exon 15, leading to transcript NM_001128147.2(NF1) (neurofibromin isoform 3). ev, empty vector.

analysis of the corresponding genomic DNA followed by Sanger sequencing (Figure 3c, Table 1).

Individuals 10, 11, and 12 were suspected to be mosaic (Table 1, Supporting Information: Table S1). In individual 10 we identified a c.4265C>T, p.(Ser1422Leu) substitution in 6/64 (10%) and 2/18 (10%) of the reads in the cycloheximide-treated and untreated samples respectively (Figure 4a, Table 1). The presence of the variant in DNA from peripheral blood from individual 10 was confirmed by AS-PCR (Figure 4b). The variant cosegregated with NF1 in multiple families in our cohort and functional analysis demonstrated that the p.(Ser1422Leu) substitution inactivates neurofibromin RAS GAP activity (Douben et al., in revision). For Individual 11, RNA samples from both normal skin fibroblasts and from cells cultured from a CAL macule were subjected to RNA-seq. There was no evidence for splicing abnormalities or for skewed expression. However, in RNA from the CAL-derived culture, a nonsense variant, c.3916C>T, p.(Arg1306*), was identified in 11/146 (~8%) of the reads. The variant was not identified by Sanger sequencing analysis of DNA isolated from peripheral blood even after careful re-evaluation of the data, but was confirmed in DNA derived from multiple CAL by targeted NGS (data not shown). In one CAL, a second pathogenic variant was identified, consistent with the hypothesis that the c.3916C>T, p.(Arg1306*) variant is the first-hit mutation and

underlying cause of NF1 in individual 11. In individual 12 we identified a premature stop, c.7043G>A, p.(Trp2348*), in 8 and 23% of the RNA-seq reads.

In individual 13 we did not find evidence for skewed expression, but did identify specific, noncanonical exon skipping events affecting pre-mRNA splicing around exon 25 (Table 1). Abnormal *NF1* pre-mRNA splicing in a small proportion of transcripts was confirmed by RT-PCR analysis, but the corresponding genomic DNA variant could not be identified.

In individuals 14 and 15, we did not identify evidence for pseudoexons, transposon insertions, or postzygotic changes in either *NF1* or *SPRED1*. In each case, *NF1* expression was bi-allelic indicating that these patients could be mosaic for a low-frequency *NF1* variant that is not detectable in fibroblast RNA, or may have an alternative genetic cause.

4 | DISCUSSION

In our study, RNA-seq analysis of NF1 NMI individuals yielded a high rate of pathogenic *NF1* variant detection. We identified disease-causing variants in 12/15 individuals and abnormal splicing in one individual for which we could not yet identify the corresponding DNA
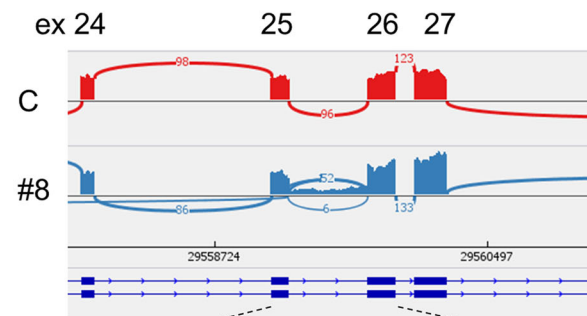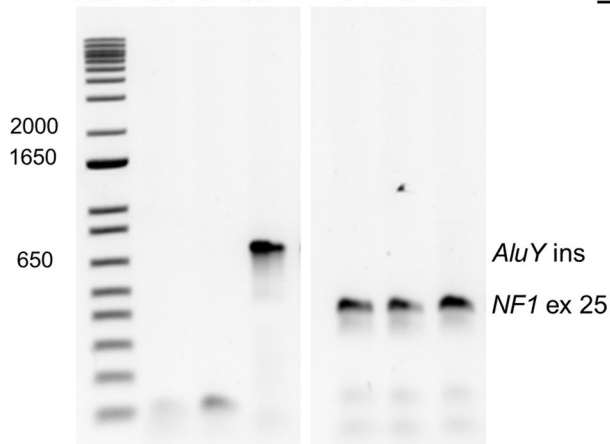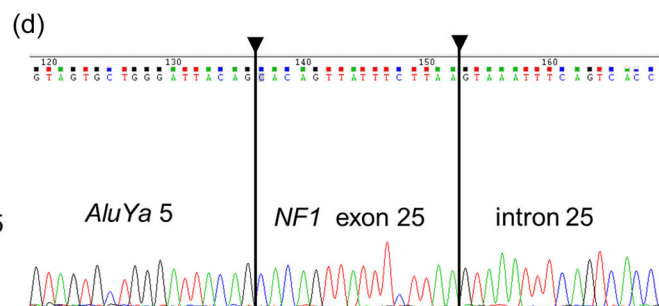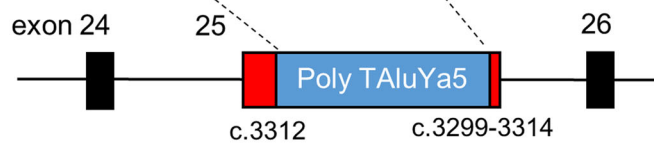
FIGURE 3 RNA-seq analysis of fibroblast mRNA identifies pathogenic transposon insertions. (a, b) Sashimi plot showing abnormal splicing around exons 25–26 in individual 8 (upper panel/a) and abnormal reads (partial exon 25 skipping, intron 25 retention). Right panel: Reads mapping from the 3′ end of exon 25 to an *AluYa5* element 600 kb upstream. (c, d) Agarose gel and electropherogram results from allele-specific PCR confirming insertion of an *AluYa5* transposon in exon 25 of *NF1* in DNA from individual 8.

abnormality. In addition, our study resulted in a direct molecular diagnosis in 17 additional affected individuals from 4 unrelated families with NF1 due to variants that are likely to result in the expression of the truncated neurofibromin isoform 3.

For cases from our clinic where other approaches fail to identify a pathogenic DNA variant, we apply RNA-seq for transcriptome-wide detection of abnormal transcripts (Cummings et al., 2017; Dekker et al., 2022; Kremer et al., 2017). We chose to analyze cultured skin fibroblast RNA as the expression of *NF1* and many other genes is high compared to expression in lymphocytes (data not shown) and because our laboratory routinely uses cultured fibroblasts for the diagnosis of metabolic disorders, enzyme deficiencies, ciliopathies and neurodevelopmental disorders (van den Bosch et al., 2011; Dekker et al., 2022; Kheradmand Kia et al., 2012; Sofou et al., 2021). Furthermore,

fibroblasts yield large amounts of high-quality RNA and can easily be stored for reculture, validation, and further research.

An established effective approach for *NF1* variant detection is cDNA-based screening of lymphocyte RNA (Ars et al., 2000; Martinez et al., 1996; Messiaen et al., 2000; Serra et al., 2001; Wimmer et al., 2011, 2020; Zatkova et al., 2004). An advantage of this RNA-first approach is that only one blood sample is required. In addition, sampling blood is typically considered to be less invasive than acquiring a skin biopsy. However, in our experience, a small skin biopsy, taken rapidly under local anesthesia, is a simple procedure, and often less stressful for the subject than drawing blood.

Because we had established a method for transcriptome-based analysis of fibroblast RNA in our diagnostic laboratory, we chose to apply this method to resolve a series of NF1 NMI cases. RNA-seq results in an unbiased overview of the transcriptional landscape,
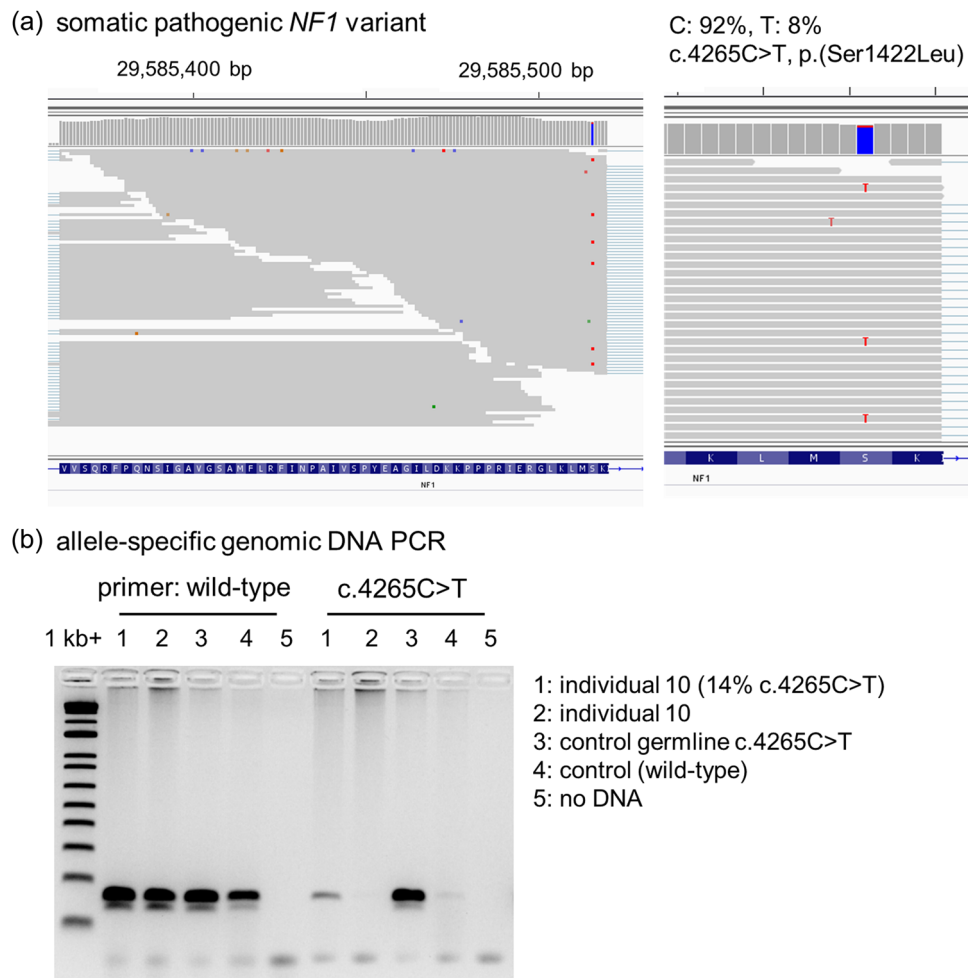
(a) somatic pathogenic *NF1* variant

29,585,400 bp   29,585,500 bp

C: 92%, T: 8%
c.4265C>T, p.(Ser1422Leu)

(b) allele-specific genomic DNA PCR

primer: wild-type   c.4265C>T

1 kb+ 1 2 3 4 5   1 2 3 4 5

1: individual 10 (14% c.4265C>T)
2: individual 10
3: control germline c.4265C>T
4: control (wild-type)
5: no DNA

FIGURE 4 RNA-seq analysis of fibroblast mRNA identifies postzygotic variants. (a) Low frequency c.4265C>T, p.(Ser1422Leu) pathogenic variant (8% reads) in individual 10. (b) Agarose gel showing products of allele-specific PCR on genomic DNA from individual 10 (c.4265C>T; p.(Ser1422Leu), DNA from blood of an individual with a germline c.4265C>T; p.(Ser1422Leu) variant was used as a positive control (sample 3).

facilitating the identification of skewed or monoallelic expression and differentiation between likely germ-line and postzygotic events. Furthermore, it is unlikely that the variants in intron 15 that resulted in the expression of the truncated transcript encoding neurofibromin isoform 3 would have been identified by cDNA screening. In this respect, other NGS-based approaches, including targeted RNA-seq analysis after enrichment for *NF1* transcripts, yielding a high gene-specific coverage, or analysis of CAL or neurofibroma tissue directly increase the capacity to detect pathogenic variants (Koster et al., 2021; Maertens et al., 2007). Nonetheless, an advantage of our approach is that it is not limited to NF1, and the same laboratory flow can be applied to most genes that are expressed at moderate to high levels in cultured fibroblasts (Dekker et al., 2022).

Pathogenic, transposable-element insertions in *NF1* have been observed previously (Wallace et al., 1991; Wimmer et al., 2011). We identified two *AluY-* insertions, in exons 25 and exon 45, the latter of which was identified in four individuals from three generations of a single family. Taken together with the identification of somatic variants, pseudoexons, and the novel finding of unconventional

isoform 3 activation, this provides further evidence for the effectiveness of RNA-seq for the molecular diagnosis of NF1 NMI individuals.

Similar to the c.1722-26T>G and c.1722-24A>G variants, the pathogenic *NF1* c.1722-11T>G variant causes skipping of exon 16 in vitro (Wimmer et al., 2020). Based on our RNA-seq data it is possible that also in this case, intron 15 retention and expression of neurofibromin isoform 3 occurs in vivo. The function of neurofibromin isoform 3 is unknown, but it lacks multiple critical functional domains and is presumably (over) expressed in individuals carrying the c.1722-26T>G, c.1722-24A>G, c.1722-14_1722-5del10, c.1722-16_1727del22, and possibly c.1722-11T>G variants.

Although visualization of the RNA-seq data in the IGV was sufficient to identify pathogenic *NF1* mRNA transcripts, analysis of patient DNA was necessary to confirm the RNA-seq data and to identify the causative DNA variant. The pseudoexons and intron 15 retention stood out immediately when visualizing the data in the IGV. In contrast, the 2 *AluY* insertions were only identified upon follow-up of a limited number of abnormal reads.

Similarly, postzygotic changes were only identified after close inspection of the mapped reads.

Although targeted NGS on DNA may increase the detection rate of (likely) pathogenic NF1 variants, our RNA-based approach is able to identify the effects of changes such as transposon insertions and deep intronic variants, on NF1 expression. Our work shows that complementary analysis of both DNA and RNA has added value, especially with the utility of skin-fibroblast transcriptome analysis to obtain a molecular diagnosis in patients meeting diagnostic criteria for NF1. Importantly, the intronic variants identified herein (Table 1) are good candidates for therapeutic antisense oligonucleotide strategies (Hammond et al., 2021).

In conclusion, RNA analysis boosts detection rates of NF1 pathogenic variants. By applying straight forward transcriptome analysis on fibroblast mRNA, we identified a diverse set of pathogenic NF1 variants consisting of deep intronic changes, transposable element insertions and postzygotic pathogenic variants in NF1 NMI individuals. Our results confirm that NF1 mRNA expression analysis in fibroblasts provides a highly complementary approach to DNA-based screening. Importantly, improved NF1 mutation detection will lead to more appropriate care for NF1 patients and their families, and provides options for prenatal testing for family planning purposes. We predict that this approach will also be useful for resolving NMI cases for other genetic disorders.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The novel variants have been deposited in LOVD3: https://databases.lovd.nl/shared/genes/NF1. Primer-sequences are available on request. The data that support the findings of this study are available from the corresponding author, with the exception of primary patient sequencing data, as they are derived from patient samples with unique variants that are impossible to guarantee anonymity for. Our institutional guidelines do not allow sharing these raw exome or genome sequencing data, as this is not part of the patient consent procedure.

ORCID

Mark Nellist http://orcid.org/0000-0003-0949-1566
Ingrid van de Laar http://orcid.org/0000-0002-2523-1230
Tjakko J. van Ham http://orcid.org/0000-0002-2175-8713

REFERENCES

Ars, E. (2000). Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. Human Molecular Genetics, 9(2), 237–247. https://doi.org/10.1093/hmg/9.2.237

Ballester, R., Marchuk, D., Boguski, M., Saulino, A., Letcher, R., Wigler, M., & Collins, F. (1990). The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. Cell, 63(4), 851–859. https://doi.org/10.1016/0092-8674(90)90151-4

van den Bosch, J., Oemardien, L. F., Srebniak, M. I., Piraud, M., Huijmans, J. G. M., Verheijen, F. W., & Ruijter, G. J. G. (2011). Prenatal screening of sialic acid storage disease and confirmation in cultured fibroblasts by LC-MS/MS. Journal of Inherited Metabolic Disease, 34(5), 1069–1073. https://doi.org/10.1007/s10545-011-9351-3

Church, D. M., Stotler, C. J., Rutter, J. L., Murrell, J. R., Trofatter, J. A., & Buckler, A. J. (1994). Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. Nature Genetics, 6(1), 98–105. https://doi.org/10.1038/ng0194-98

Colman, S. D., Rasmussen, S. A., Ho, V. T., Abernathy, C. R., & Wallace, M. R. (1996). Somatic mosaicism in a patient with neurofibromatosis type 1. American Journal of Human Genetics, 58(3), 484–490. https://www.ncbi.nlm.nih.gov/pubmed/8644707

Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., Bolduc, V., Waddell, L. B., Sandaradura, S. A., O'Grady, G. L., Estrella, E., Reddy, H. M., Zhao, F., Weisburd, B., Karczewski, K. J., O'Donnell-Luria, A. H., Birnbaum, D., Sarkozy, A., Hu, Y., ... MacArthur, D. G. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Science Translational Medicine, 9(386), eaal5209. https://doi.org/10.1126/scitranslmed.aal5209

Dekker, J., Schot, R., Bongaerts, M., de Valk, W. G., van Veghel-Plandsoen, M. M., Monfils, K., Douben, H., Elfferich, P., Kasteleijn, E., van Unen, L. M. A., Geeven, G., Saris, J. J., van Ierland, Y., Verheijen, F. W., van der Sterre, M. L. T., Niaraki, F. S., Huidekoper, H. H., Williams, M., Wilke, M., ... van Ham, T. J. (2022). RNA-sequencing improves diagnosis for neurodevelopmental disorders by identifying pathogenic non-coding variants and reinterpretation of coding variants. medRxiv, 1–40. https://doi.org/10.1101/2022.06.05.22275956

Dufner-Almeida, L. G., Nanhoe, S., Zonta, A., Hosseinzadeh, M., Kom-Gortat, R., Elfferich, P., Schaaf, G., Kenter, A., Kümmel, D., Migone, N., Povey, S., Ekong, R., & Nellist, M. (2020). Comparison of the functional and structural characteristics of rare TSC2 variants with clinical and genetic findings. Human Mutation, 41(4), 759–773. https://doi.org/10.1002/humu.23963

Evans, D. G., Bowers, N., Burkitt-Wright, E., Miles, E., Garg, S., Scott-Kitching, V., Penman-Splitt, M., Dobbie, A., Howard, E., Ealing, J., Vassalo, G., Wallace, A. J., Newman, W., & Huson, S. M. (2016). Comprehensive RNA analysis of the NF1 gene in classically affected NF1 affected individuals meeting NIH criteria has high sensitivity and mutation negative testing is reassuring in isolated cases with pigmentary features only. EBioMedicine, 7, 212–220. https://doi.org/10.1016/j.ebiom.2016.04.005

gnomAD v2.1.1. https://gnomad.broadinstitute.org/

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison 3rd, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods, 6(5), 343–345. https://doi.org/10.1038/nmeth

Hammond, S. M., Aartsma-Rus, A., Alves, S., Borgos, S. E., Buijsen, R. A. M., Collin, R. W. J., Covello, G., Denti, M. A., Desviat, L. R., Echevarría, L., Foged, C., Gaina, G., Garanto, A., Goyenvalle, A. T., Guzowska, M., Holodnuka, I., Jones, D. R., Krause, S., Lehto, T., ... Arechavala-Gomeza, V. (2021). Delivery of oligonucleotide-based therapeutics: Challenges and opportunities. EMBO Molecular Medicine, 13(4), e13243. https://doi.org/10.15252/emmm.202013243

Jett, K., & Friedman, J. M. (2010). Clinical and genetic aspects of neurofibromatosis 1. Genetics in Medicine, 12(1), 1–11. https://doi.org/10.1097/GIM.0b013e3181bf15e3

Kheradmand Kia, S., Verbeek, E., Engelen, E., Schot, R., Poot, R. A., de Coo, I. F. M., Lequin, M. H., Poulton, C. J., Pourfarzad, F., Grosveld, F. G., Brehm, A., de Wit, M. C. Y., Oegema, R., Dobyns, W. B., Verheijen, F. W., & Mancini, G. M. S. (2012). RTTN mutations link primary cilia function to organization of the human cerebral cortex. *The American Journal of Human Genetics*, *91*(3), 533–540. https://doi.org/10.1016/j.ajhg.2012.07.008

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Koster, R., Brandão, R. D., Tserpelis, D., van Roozendaal, C. E. P., van Oosterhoud, C. N., Claes, K. B. M., Paulussen, A. D. C., Sinnema, M., Vreeburg, M., van der Schoot, V., Stumpel, C. T. R. M., Broen, M. P. G., Spruijt, L., Jongmans, M. C. J., Lesnik Oberstein, S. A. J., Plomp, A. S., Misra-Isrie, M., Duijkers, F. A., Louwers, M. J., ... Blok, M. J. (2021). Pathogenic neurofibromatosis type 1 (NF1) RNA splicing resolved by targeted RNAseq. *NPJ Genomic Medicine*, *6*(1), 95. https://doi.org/10.1038/s41525-021-00258-w

Kremer, L. S., Bader, D. M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A., Haack, T. B., Graf, E., Schwarzmayr, T., Terrile, C., Koňaříková, E., Repp, B., Kastenmüller, G., Adamski, J., Lichtner, P., Leonhardt, C., Funalot, B., Donati, A., Tiranti, V., ... Prokisch, H. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nature Communications*, *8*, 15824. https://doi.org/10.1038/ncomms15824

Lazaro, C., Ravella, A., Gaona, A., Volpini, V., & Estivill, X. (1994). Neurofibromatosis type 1 due to germ-line mosaicism in a clinically normal father. *New England Journal of Medicine*, *331*(21), 1403-1407. https://doi.org/10.1056/NEJM199411243312102

Legius, E., Messiaen, L., Wolkenstein, P., Pancza, P., Avery, R. A., Berman, Y., Blakeley, J., Babovic-Vuksanovic, D., Cunha, K. S., Ferner, R., Fisher, M. J., Friedman, J. M., Gutmann, D. H., Kehrer-Sawatzki, H., Korf, B. R., Mautner, V. F., Peltonen, S., Rauen, K. A., Riccardi, V., ... Plotkin S. R. (2021). Revised diagnostic criteria for neurofibromatosis type 1 and Legius syndrome: An international consensus recommendation. *Genetics in Medicine*, *23*(8), 1506–1513. https://doi.org/10.1038/s41436-021-01170-5

Maertens, O., De Schepper, S., Vandesompele, J., Brems, H., Heyns, I., Janssens, S., Speleman, F., Legius, E., & Messiaen, L. (2007). Molecular dissection of isolated disease features in mosaic neuro-fibromatosis type 1. *The American Journal of Human Genetics*, *81*(2), 243–251. https://doi.org/10.1086/519562

Martinez, J. M., Breidenbach, H. H., & Cawthon, R. (1996). Long RT-PCR of the entire 8.5-kb NF1 open reading frame and mutation detection on agarose gels. *Genome Research*, *6*(1), 58–66. https://doi.org/10.1101/gr.6.1.58

Messiaen, L. M., Callens, T., Mortier, G., Beysen, D., Vandenbroucke, I., Van Roy, N., Speleman, F., & Paepe, A. D. (2000). Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Human Mutation*, *15*(6), 541–555. https://doi.org/10.1002/1098-1004(200006)15:6.3.0.CO;2-N

van Minkelen, R., van Bever, Y., Kromosoeto, J. N. R., Withagen-Hermans, C. J., Nieuwlaat, A., Halley, D. J. J., & van den Ouweland, A. M. W. (2014). A clinical and genetic overview of 18 years neuro-fibromatosis type 1 molecular diagnostics in the Netherlands. *Clinical Genetics*, *85*(4), 318-327. https://doi.org/10.1111/cge.12187

Pasmant, E., Parfait, B., Luscan, A., Goussard, P., Briand-Suleau, A., Laurendeau, I., Fouveaut, C., Leroy, C., Montadert, A., Wolkenstein, P., Vidaud, M., & Vidaud, D. (2015). Neurofibromatosis type 1 molecular diagnosis: What can NGS do for you when you have a large gene with loss of function mutations? *European Journal of Human Genetics*, *23*(5), 596–601. https://doi.org/10.1038/ejhg.2014.145

Primer 3 DNA primer design application. https://primer3.ut.ee/

Pros, E., Gómez, C., Martín, T., Fábregas, P., Serra, E., & Lázaro, C. (2008). Nature and mRNA effect of 282 different NF1 point mutations: Focus on splicing alterations. *Human Mutation*, *29*(9), E173-E193. https://doi.org/10.1002/humu.20826

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & ACMG Laboratory Quality Assurance Committee. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Sabbagh, A., Pasmant, E., Imbard, A., Luscan, A., Soares, M., Blanché, H., Laurendeau, I., Ferkal, S., Vidaud, M., Pinson, S., Bellanné-Chantelot, C., Vidaud, D., Parfait, B., & Wolkenstein, P. (2013). NF1 molecular characterization and neurofibromatosis type I genotype-phenotype correlation: The French experience. *Human Mutation*, *34*(11), 1510–1518. https://doi.org/10.1002/humu.22392

Serra, E., Ars, E., Ravella, A., Sánchez, A., Puig, S, Rosenbaum, T., Estivill, X., & Lázaro, C (2001). Somatic NF1 mutational spectrum in benign neurofibromas: MRNA splice defects are common among point mutations. *Human Genetics*, *108*(5), 416–429. https://doi.org/10.1007/s004390100514

Sofou, K., Meier, K., Sanderson, L. E., Kaminski, D., Montoliu-Gaya, L., Samuelsson, E., Blomqvist, M., Agholme, L., Gärtner, J., Mühlhausen, C., Darin, N., Barakat, T. S., Schlotawa, L., Ham, T., Asin Cayuela, J., & Sterky, F. H. (2021). Bi-allelic VPS16 variants limit HOPS/CORVET levels and cause a mucopolysaccharidosis-like disease. *EMBO Molecular Medicine*, *13*(5), e13376. https://doi.org/10.15252/emmm.202013376

Suzuki, H., Takahashi, K., Kubota, Y., & Shibahara, S. (1992). Molecular cloning of a cDNA coding for neurofibromatosis type 1 protein isoform lacking the domain related to ras GTPase-activating protein. *Biochemical and Biophysical Research Communications*, *187*(2), 984–990. https://doi.org/10.1016/0006-291x(92)91294-z

Trovó-Marqui, A., & Tajara, E. (2006). Neurofibromin: A general outlook. *Clinical Genetics*, *70*(1), 1–13. https://doi.org/10.1111/j.1399-0004.2006.00639.x

Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W., & Collins, F. S. (1991). A de novo Alu insertion results in neurofibromatosis type 1. *Nature*, *353*(6347), 864–866. https://doi.org/10.1038/353864a0

Wang, L., Nie, J., Sicotte, H., Li, Y., Eckel-Passow, J. E., Dasari, S., Vedell, P. T., Barman, P., Wang, L., Weinshiboum, R., Jen, J., Huang, H., Kohli, M., & Kocher, J. P. A. (2016). Measure transcript integrity using RNA-seq data. *BMC Bioinformatics*, *17*, 58. https://doi.org/10.1186/s12859-016-0922-z

Wimmer, K., Callens, T., Wernstedt, A., & Messiaen, L. (2011). The *NF1* gene contains hotspots for *L1* endonuclease-dependent de novo insertion. *PLoS Genetics*, *7*(11), e1002371. https://doi.org/10.1371/journal.pgen.1002371

Wimmer, K., Schamschula, E., Wernstedt, A., Traunfellner, P., Amberger, A., Zschocke, J., Kroisel, P., Chen, Y., Callens, T., & Messiaen, L. (2020). AG-exclusion zone revisited: Lessons to learn from 91 intronic NF1 3′ splice site mutations outside the canonical

AG-dinucleotides. *Human Mutation*, *41*(6), 1145–1156. https://doi.org/10.1002/humu.24005

Zatkova, A., Messiaen, L., Vandenbroucke, I., Wieser, R., Fonatsch, C., Krainer, A. R., & Wimmer, K. (2004). Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Human Mutation*, *24*(6), 491–501. https://doi.org/10.1002/humu.20103

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Douben, H. C. W., Nellist, M., van Unen, L., Elfferich, P., Kasteleijn, E., Hoogeveen-Westerveld, M., Louwen, J., van Veghel-Plandsoen, M., de Valk, W., Saris, J. J., Hendriks, F., Korpershoek, E., Hoefsloot, L. H., van Vliet, M., van Bever, Y., van de Laar, I., Aten, E., Lachmeijer, A. M. A., Taal, W., … van Ham, T. J. (2022). High-yield identification of pathogenic *NF1* variants by skin fibroblast transcriptome screening after apparently normal diagnostic DNA testing. *Human Mutation*, 1–11. https://doi.org/10.1002/humu.24487