RESEARCH ARTICLE

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

Hannie C. W. Douben1 | Mark Nellist1 | Leontine van Unen1 | Peter Elfferich1 | Esmee Kasteleijn1 | Marianne Hoogeveen-Westerveld1 | Jesse Louwen1 | Monique van Veghel-Plandsoen1 | Walter de Valk1 | Jasper J. Saris1 | Femke Hendriks1 | Esther Korpershoek1,2 | Lies H. Hofsloot1 | Margreethe van Vliet1,3 | Yolande van Bever1 | Ingrid van de Laar1 | Emmelien Aten4 | Augusta M. A. Lachmeijer5 | Walter Taal3,6 | Lisa van den Bersselaar1 | Juliette Schuermans5 | Rianne Oostenbrink3,7 | Rick van Minkelen1,3 | Yvette van Ierland1,3 | Tjakko J. van Ham1 ♦

1Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
2Department of Pathology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
3ENCORE Expertise Center for Neurodevelopmental Disorders, Erasmus MC, Rotterdam, The Netherlands
4Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
5Department of Genetics, Division Laboratories, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands
6Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands
7Department 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

Funding information
Erasmus University Rotterdam Fellowship

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

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. Human Mutation published by Wiley Periodicals LLC.
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.

**KEYWORDS**
exon skipping, molecular diagnostics, mRNA splicing, neurofibromatosis type 1, noncoding variants, RNA-sequencing

## 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; Pasman 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 pre-mRNA splicing and/or are present at a low frequency due to mosaicism (Ares et al., 2000; Evans et al., 2016; Koster et al., 2021; Martínez 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 2a–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 non-canonical 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 2a–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.1722−267>T>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

<table>
<thead>
<tr>
<th>Patient</th>
<th>NF1 variant NC_000017.10(NM_000267.3) (position)</th>
<th>Effect on transcript</th>
<th>Type mutation/effect</th>
<th>Validation</th>
<th>Detected previously/freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c.1527 + 1187C&gt;G (i13)</td>
<td>r.1527_1528ins[1527+1104_1527+1182]</td>
<td>Creates 5’ donor; Pseudo-exon, skewed expr.</td>
<td>PCR</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>c.888 + 789A&gt;G (i8)</td>
<td>r.888,889ins[888+656,888+784], r.888,889ins[888+710,888+784]</td>
<td>Strengthens noncanonical 5’ donor; Pseudo-exon, skewed expr.</td>
<td>PCR</td>
<td>ClinVar: 664551 (Evans et al., 2016)</td>
</tr>
<tr>
<td>3</td>
<td>c.888 + 789A&gt;G (i8)</td>
<td>r.888ins888 + 656,888 + 784/ r.888ins888 + 710,888 + 784</td>
<td>Strengthens noncanonical 5’ donor; Pseudo-exon, skewed expr.</td>
<td>PCR</td>
<td>ClinVar: 664551 (Evans et al., 2016)</td>
</tr>
<tr>
<td>5</td>
<td>c.5749 + 332A&gt;G (i39)</td>
<td>r.5749ins5749 + 156,5949 + 331, p.(Ser1917Lysfs*3)</td>
<td>Creates 5’ donor; Pseudo-exon</td>
<td>PCR</td>
<td>ClinVar: 216065</td>
</tr>
<tr>
<td>6</td>
<td>c.5749 + 267A&gt;G (i39)</td>
<td>r.5749ins5749 + 156,5949 + 267, p.(Ser1917Lysfs*3)</td>
<td>Creates 5’ donor; Pseudo-exon</td>
<td>PCR</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>c.1722 - 26T&gt;G (i15)</td>
<td>r.1722_1845del, p.(Ser575Argfs*15)</td>
<td>Branch point/activation isoform 3, skewed expr.</td>
<td>exon trapping</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>c.3313_3313ins[AluYα5;3299_3312dup] (e25)</td>
<td>skipping e23-25: r.2991_3314del, r.3275_3314del, r.3211_3314del</td>
<td>Transposon ins. (AluYα5), skewed expr.</td>
<td>PCR</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>c.6852_6853ins[T?;AluYα6839_6852dup] (e46)</td>
<td>r.6757_6858del, p.(Ala2253_Lys2286del)</td>
<td>Transposon ins. (AluY), skewed expr.</td>
<td>PCR</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>c.4265C&gt;T, p.(ser1422Leu) (e32)</td>
<td>r.4265c&gt;u</td>
<td>Missense, bi-allelic expr.</td>
<td>AS-PCR</td>
<td>Class 5 ClinVar: 527535</td>
</tr>
<tr>
<td>11</td>
<td>c.3916C&gt;T, p.(Arg1306*) (e29)</td>
<td>r.3916c&gt;u</td>
<td>Stop gain, bi-allelic expr.</td>
<td>AS-PCR</td>
<td>Class 5 ClinVar: 404592</td>
</tr>
<tr>
<td>12</td>
<td>c.7043G&gt;A, p.(Trp2348*) (e48)</td>
<td>Premature stop</td>
<td>Premature stop</td>
<td>AS-PCR</td>
<td>Class 5 ClinVar: 996432</td>
</tr>
<tr>
<td>13</td>
<td>Skipping exon 25, NMI (e25)</td>
<td>confirmed low fraction abnormal splicing</td>
<td>Possibly postzygotic, bi-allelic expr.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>14</td>
<td>NMI</td>
<td>n.a.</td>
<td>Possibly postzygotic, bi-allelic expr.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>15</td>
<td>NMI</td>
<td>n.a.</td>
<td>Possibly postzygotic, bi-allelic expr.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

Abbreviations: AS-PCR, allele-specific PCR; e/i, exon/intron.
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.
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.3916G>A, p.(Trp2348*), 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.3916G>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
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; Kreme 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,
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.

**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).
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.

ACKNOWLEDGMENTS
TvH was supported by an Erasmus University Rotterdam (EUR) fellowship.

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


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