Genetics in Medicine (2023) 25, 100927

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

De novo *PHF5A* variants are associated with craniofacial abnormalities, developmental delay, and hypospadias

Frederike L. Harms[1], Alexander J.M. Dingemans[2], Maja Hempel[1,3], Rolph Pfundt[2], Tatjana Bierhals[1], Christian Casar[4], Christian Müller[4], Jikke-Mien F. Niermeijer[5], Jan Fischer[6], Arne Jahn[6], Christoph Hübner[7], Silvia Majore[8], Emanuele Agolini[9], Antonio Novelli[9], Jasper van der Smagt[10], Robert Ernst[10], Ellen van Binsbergen[10], Grazia M.S. Mancini[11], Marjon van Slegtenhorst[11], Tahsin Stefan Barakat[11,12], Emma L. Wakeling[13], Arveen Kamath[14], Lilian Downie[15,16], Lynn Pais[17], Susan M. White[15,16], Bert B.A. de Vries[2,*], Kerstin Kutsche[1,*]

ARTICLE INFO

Article history:
Received 29 December 2022
Received in revised form
30 June 2023
Accepted 2 July 2023
Available online 6 July 2023

Keywords:
Craniofacial spliceosomopathies
Exome
Loss of function
Nager syndrome
Negative autoregulation

ABSTRACT

Purpose: The SF3B splicing complex is composed of SF3B1-6 and PHF5A. We report a developmental disorder caused by de novo variants in *PHF5A*.
Methods: Clinical, genomic, and functional studies using subject-derived fibroblasts and a heterologous cellular system were performed.
Results: We studied 9 subjects with congenital malformations, including preauricular tags and hypospadias, growth abnormalities, and developmental delay who had de novo heterozygous *PHF5A* variants, including 4 loss-of-function (LOF), 3 missense, 1 splice, and 1 start-loss variant. In subject-derived fibroblasts with *PHF5A* LOF variants, wild-type and variant *PHF5A* mRNAs had a 1:1 ratio, and *PHF5A* mRNA levels were normal. Transcriptome sequencing revealed alternative promoter use and downregulated genes involved in cell-cycle regulation. Subject and control fibroblasts had similar amounts of PHF5A with the predicted wild-type molecular weight and of SF3B1-3 and SF3B6. SF3B complex formation was unaffected in 2 subject cell lines.
Conclusion: Our data suggest the existence of feedback mechanisms in fibroblasts with *PHF5A* LOF variants to maintain normal levels of SF3B components. These compensatory mechanisms in subject fibroblasts with *PHF5A* or *SF3B4* LOF variants suggest disturbed autoregulation of mutated splicing factor genes in specific cell types, that is, neural crest cells, during embryonic development rather than haploinsufficiency as pathomechanism.

Bert B.A. de Vries and Kerstin Kutsche contributed equally.

*Correspondence and requests for materials should be addressed to Kerstin Kutsche, Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. *Email address:* kkutsche@uke.de OR Bert B.A. de Vries, Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands. *Email address:* Bert.deVries@radboudumc.nl

Affiliations are at the end of the document.

doi: https://doi.org/10.1016/j.gim.2023.100927
1098-3600/

Introduction

Splicing is an essential process in eukaryotes whereby the spliceosome removes the intron sequences and ligates the exon sequences together from pre-mRNAs to produce mature and functional mRNAs. The spliceosomal complex is a dynamic macromolecular machine comprising multiple uridine-rich (U1, U2, U4/U6, and U5) small nuclear RNAs and associated ribonucleoproteins (snRNPs), as well as other proteins associated with these snRNPs.[1-3] The pre-mRNA contains conserved *cis*-splicing consensus sequences that include the 5′ (GT) and 3′ (AG) splice sites, a polypyrimidine tract, and the branch point sequences, which surround the branch point adenosine. These conserved intronic sequences are required for the recognition and assembly of the spliceosome, which undergoes conformational rearrangements to mediate the splicing process.[1,3]

One of the numerous snRNPs choreographing splicing is the U2 snRNP complex, which assembles in an adenosine triphosphate-dependent manner at the 3′ acceptor site of the intron. The U2 snRNP contains stable core subunits and the dynamically associated sub-complexes SF3A and SF3B. The SF3B complex is mainly involved in recognition of the branch point adenosine.[4] SF3B contains 7 proteins, namely, SF3B1, SF3B2, SF3B3, SF3B4, SF3B5, SF3B6, and PHF5A.[5-7] PHF5A, a core component of the SF3B complex, is a ubiquitously expressed and evolutionarily highly conserved nuclear protein.[8] The protein is small and globular, has a molecular mass of about 12 kDa, and forms a unique trefoil knot, which is stabilized by 3 zinc-finger motifs. There are no close homologs of PHF5A in the human genome, suggesting an important biological role for this protein. It has been confirmed that PHF5A, together with SF3B1, forms the branch point adenosine-binding pocket, indicating an essential role of PHF5A in splicing, not only in humans but also in plants.[4,9,10]

Pathogenic variants in genes encoding core components of the spliceosome cause a spectrum of disorders, named spliceosomopathies. This group of diseases can be subdivided into craniofacial and myelodysplastic syndromes (MDS)-related spliceosomopathies.[11] Recurrent somatic pathogenic variants in 4 genes (including *SF3B1* encoding 1 SF3B complex component) were found in myelodysplastic syndromes and other myeloid neoplasms (OMIM 614286).[11,12] Individuals affected by craniofacial spliceosomopathies can show microcephaly, micrognathia, malar hypoplasia, external ear anomalies, ophthalmological anomalies, and skeletal and/or heart defects. The 9 currently associated disease genes are *SNRPB*, *RNU4ATAC*, *SF3B2*, *SF3B4*, *PUF60*, *EFTUD2*, *TXNL4*, *EIF4A3*, and *CWC27*.[11,13] Heterozygous pathogenic variants in *SNRPB* cause the cerebrocostomandibular syndrome (OMIM 117650).[14] Biallelic *RNU4ATAC* pathogenic variants are associated with a phenotypic spectrum called RNU4atac-opathy, including microcephalic osteodysplastic primordial dwarfism type I/III, Roifman syndrome, and Lowry-Wood syndrome (OMIM 210710, 226960, and 616651).[15-19] Verheij syndrome is caused by heterozygous de novo interstitial deletions of the 8q24.3 region or single-nucleotide pathogenic variants in *PUF60* (OMIM 615583).[20,21] Heterozygous *EFTUD2* pathogenic variants underlie mandibulofacial dysostosis with microcephaly (OMIM 610536).[22,23] *TXNL4A* and *EIF4A3* are the disease genes for autosomal recessive Burn-McKeown syndrome (OMIM 608572) and Robin sequence with cleft mandible and limb anomalies (OMIM 268305), respectively[24-26] and biallelic *CWC27* pathogenic variants cause retinitis pigmentosa with or without skeletal anomalies (OMIM 250410).[27] Developmental delay and intellectual disability are characteristic clinical features of Verheij syndrome and mandibulofacial dysostosis with microcephaly,[22,28] whereas individuals with *TXNL4A*-related craniofacial disorders generally have normal cognitive functioning.[25]

SF3B2 and SF3B4 belong to the SF3B complex.[5,6] Heterozygous loss-of-function (LOF) variants in *SF3B4* cause Nager syndrome and Rodriguez syndrome, with Rodriguez syndrome representing the more severe end of the clinical spectrum (OMIM 154400).[29,30] Affected individuals show down slanted palpebral fissures, malar hypoplasia, external ear anomalies, and micrognathia as craniofacial defects and preaxial and/or postaxial limb abnormalities. Individuals with Rodriguez syndrome often die before or around birth from respiratory complications.[29-36] Heterozygous LOF variants in *SF3B2* have recently been reported in 7 families with craniofacial macrosomia (OMIM 164210). Subjects have craniofacial abnormalities, such as mandibular hypoplasia, microtia, facial and preauricular tags, and skeletal abnormalities, such as cervical ribs.[37] Although several genes coding for components of the SF3B complex are associated with genetic disorders, evidence for a disease phenotype associated with pathogenic variants in *PHF5A* is limited to a single case report. A de novo nonsense variant c.162C>A / p.(Tyr54*) in *PHF5A* has recently been reported in a girl with left microtia and bilateral absence of 12th ribs.[38] In this study, we report on 9 subjects with de novo heterozygous *PHF5A* variants. We describe functional data using fibroblasts from 5 affected individuals, further strengthening the association of *PHF5A* with a human disease.

Materials and Methods

Subjects

The clinical data were converted to human phenotype ontology terms[39] to ensure a standardized data set and the complete phenotype of included individuals is available in the Supplemental Material of this paper in the Phenopacket format.[40] To investigate whether subjects with a de novo variant in *PHF5A* have a distinguishable phenotype, the clinical data were analyzed using PhenoScore (https://www.medrxiv.org/content/10.1101/2022.10.24.22281480v1). We

performed 2 analyses: one did not include the 3 subjects with a *PHF5A* missense variant because we classified the variants as variants of uncertain significance (see below), whereas a second analysis included all 9 subjects described in this study. PhenoScore is a tool that enables the quantification of phenotypic similarity between individuals. In this case, we used PhenoScore to determine whether the clinical features are different from those of control individuals. These controls are individuals with other neurodevelopmental disorders seen at our outpatient clinic, matched with the subjects with variants in *PHF5A* on age, sex, and ethnicity. By training a machine-learning model on the phenotypic data and performing a permutation test, we can determine whether the phenotype is actually different than that of control individuals (for the complete details of the analysis, see https://www.medrxiv.org/content/10.1101/2022.10.24.22281480v1).

Exome and genome sequencing

Trio exome sequencing was performed in subjects 1 to 8 and parents and trio genome sequencing was performed in subject 9 and parents.

Technical details and information on exome sequencing and genome sequencing are described in the Supplemental Methods.

Variant validation

Sequence validation of all candidate variants in leukocyte-derived DNA from subjects 3 and 4 and parents, and confirmation of the *PHF5A* variants in fibroblast-derived DNA from subjects 1 to 4 and 9, were performed by Sanger-sequencing (Supplemental Figure 1). Sequences of primers designed to amplify selected coding exons of *PHF5A* and exon-intron boundaries (NM_032758.4) are described in Supplemental Table 1. Amplicons were directly sequenced using the ABI BigDye Terminator Sequencing kit (Applied Biosystems) and an automated capillary sequencer (ABI 3500, Applied Biosystems). Sequence electropherograms were analyzed using the Sequence Pilot software (JSI Medical Systems).

Transcriptome sequencing and data analysis

Total RNA was extracted (Monarch Total RNA Miniprep Kit, New England Biolabs) from cultured primary fibroblasts of subjects 1, 2, 4, and 9. RNA concentration and purity of the samples were assessed by the Epoch Microplate Spectrophotometer (BioTek). RNA sequencing was done using polyA-enriched RNA sequencing (TruSeq Stranded mRNA Kit, Illumina). Paired-end sequencing (2 × 100 bp) was performed with an Illumina NovaSeq 6000 instrument at a sequencing depth of 60 to 100 million paired reads per sample and a minimum output of 12 Gb per sample. Output bcl files were converted to fastq files and demultiplexed using bcl2fastq (v2.20, Illumina). Adapter trimming was performed with fastp (v0.23.2).[41] Sequence reads were aligned to the human reference genome assembly (GRCh38.106) using STAR (v2.7.10b)[42] in 2-pass mode. Of note, RNA from subject 2 was sequenced on another sequencing run compared with RNA of subjects 1, 4, and 9. For all subsequent analyses, data from 42 previously sequenced fibroblast-derived RNA samples were used as controls.[43]

Global splicing events were counted for each sample based on the proportion-spliced-in values (PSI) determined by SUPPA2.[44] Read trimming was performed with fastp[41] and transcripts were quantified using Salmon[45] based on human genome assembly GRCh38, GENCODE version 40. Different splice event types were generated by running the SUPPA2 method *generateEvents* on the GENCODE transcript annotation file. The SUPPA2 method *psiPerEvent* was applied to quantify the PSI for each event type and sample. Events with a PSI $\geq$ 20% were considered to be spliced in and were counted.

For detection and quantification of individual alternative splicing events, LeafCutter v0.2.9 was used with default parameters on the splice junctions generated in the STAR 2-pass alignment.[46] For differential intron excision analysis, the minimum number of samples supporting an intron and the minimum number of samples per group were reduced to 2. Differentially spliced clusters with a false discovery rate of $\leq$0.05 and a maximal absolute delta PSI $\geq$ 0.3 were visualized as Sashimi plots.

Detection of aberrantly expressed genes was performed with OUTRIDER (v1.4.2).[47] Lowly expressed genes (fragments per kilobase per millions of reads $\leq$ 1 observed in at least 95% of all samples) were removed prior analysis. After filtering, 12,326 genes remained. OUTRIDER's heuristic to automatically determine the autoencoder's optimal encoding dimension suggested a value of 6 for our data. Accordingly, this value was applied. Significance was determined by adjusted *P* value of .1.

Differential expression was assessed using DESeq2 (v1.34.0).[48] A gene was considered significantly differentially expressed if the corresponding absolute log2-transformed fold change (log2FC) was not less than 1 and a false discovery rate cutoff of 0.1.

Cell culture and transfection of HEK293T cells

Primary dermal fibroblasts were obtained from a skin biopsy of subjects 1 to 4 and 9 and 3 female (4 years old) and 7 male healthy control individuals (male control [mCtrl.] 1 [3 years], mCtrl. 2 [7 years], mCtrl. 3 [9 years], mCtrl. 4 [16 years], mCtrl. 5 [5 years], mCtrl. 6 [34 years], and mCtrl. 7 [52 years]). All skin biopsies were taken from unaffected skin in the 5 subjects: from the upper arm in subjects 1 and 3, from the forearm in subject 2, from the buttock area in subject 4, and from the back in subject 9. Fibroblasts and

HEK293T cells were cultured in Dulbecco's modified Eagle medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (GE Healthcare) and penicillin-streptomycin (100 U/mL and 100 mg/mL, respectively; Thermo Fisher Scientific). For all experiments, the same passage number of subject and control fibroblasts was used. Primary fibroblasts were regularly tested for mycoplasma contamination and confirmed to be mycoplasma free. HEK293T cells were transiently transfected with Jet-Optimus (Polyplus Transfection) according to the manufacturer's protocol and subsequently cultured in 10% Dulbecco's modified Eagle medium overnight.

RNA isolation, cDNA synthesis, and quantitative reverse transcription PCR (RT-qPCR)

Total RNA was extracted from cultured primary fibroblasts of subjects and controls (Monarch Total RNA Miniprep kit, New England BioLabs). Leukocyte-derived RNA from PAXgene blood RNA tube of subject 4 was isolated using the PAXgene Blood RNA Kit IVD (Qiagen). RNA concentration and purity of the samples were assessed by use of the Microplate Spectrophotometer Epoch (BioTek). 1 microgram of total RNA were reverse transcribed (Luna-Script RT Super Mix kit, New England BioLabs). Technical triplicates of quantitative reverse transcription polymerase chain reaction (RT-qPCR) samples were prepared as a 10 µL approach with the SYBR Green I-based Luna Universal qPCR Master Mix (New England BioLabs), 500 nM of each primer, and 1 µL of the reverse transcription reaction. Primer sequences for RT-qPCR are described in Supplemental Table 1. RT-qPCR was performed using the QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) equipped with QuantStudio Design&Analysis Software v1.4.3 (Thermo Fisher Scientific). The PCR conditions included a pre-run at 95 °C for 5 minutes, followed by 40 cycles of 30 seconds at 95 °C, 30 seconds at 58 °C, and 45 seconds at 72 °C. PCR amplification specificity was determined by melting curve analysis with a range from 60 °C to 95 °C. The values of the cycle threshold of the target mRNAs were normalized to the housekeeping mRNA of *GAPDH*. For relative gene expression the comparative cycle threshold (ΔΔCT) values were calculated with the QuantStudio Design&Analysis Software (Thermo Fisher Scientific) with *GAPDH* as housekeeping gene and expressed as x-fold change to control 2.

Plasmid information and cloning procedures

The coding region of human wild-type *PHF5A* (NM_032758.4) was amplified by using *PHF5A*-specific primers including the FLAG tag sequence either in the forward primer sequence (N-terminal FLAG tag) or in the reverse primer sequence (C-terminal FLAG tag) using human fibroblast-derived complementary DNA (cDNA) as template. Purified PCR products were cloned between *Bgl*II and *Eco*RI restriction sites of pIRES2-EGFP vector using In-Fusion HD cloning kit (Takara) according to the manufacturer's protocol. *PHF5A* variants c.44C>T / p.(Ala15Val) (A15V) and c.185G>A / p.(Gly62Glu) (G62E) were introduced into *PHF5A* cDNA using the QuikChange II Site-Directed Mutagenesis kit (Agilent). All constructs were regularly sequenced for integrity, and primer sequences for In-Fusion cloning and site-directed mutagenesis are described in Supplemental Table 1.

Antibodies

Primary antibodies: normal mouse IgG (#12-371; Upstate), mouse monoclonal anti-Flag (Sigma-Aldrich; #F3165; clone M2; 1:1,000), hFAB Rhodamine anti-GAPDH (#12004167; Bio-Rad; 1:10,000), mouse monoclonal anti-GFP (Bio-Legend; #902601; clone B34; 1:5,000), polyclonal rabbit anti-PHF5A (#15554-1-AP; Proteintech; 1:500), monoclonal mouse anti-SF3B1/Sap155 (#D221-3; MBL International; 1:1,000), polyclonal rabbit anti-SF3B2 (#NB100-79848; Novus Biologicals; 1:1,000), polyclonal rabbit anti-SF3B3 (#14577-1-AP; Proteintech; 1:1,000), and polyclonal rabbit anti-SF3B6/SF3B13 (#12379-1-AP; Proteintech; 1:500).

Secondary antibodies: StarBright Blue 700 goat anti-mouse IgG (#12004158; BioRad; 1:10,000); StarBright Blue 700 goat anti-rabbit IgG (#12004161; BioRad; 1:10,000).

Immunoblotting

Next day after transfection of HEK293T cells, whole-cell lysates were prepared in ice-cold lysis buffer (50 mM Tris-HCl, 120 mM NaCl, 1 mM EDTA, 0.5% NP-40) supplemented with Mini Protease Inhibitor and PhosSTOP (Roche) and lysed for 10 minutes on ice. Fibroblasts were harvested in ice-cold RIPA buffer (50 mM Tris-HCl, pH 8.0; 150 mM NaCl; 1% NP-40; 0.5% sodium deoxycholate; 0.1% sodium dodecyl sulfate [SDS]) supplemented with Mini Protease Inhibitor and PhosSTOP and lysed for 10 minutes on ice. Cell debris was removed by centrifugation for 10 minutes. Protein extracts were separated on sodium dodecyl-sulfate polyacrylamide gel electrophoresis under denaturing conditions and transferred to polyvinylidene fluoride membranes. Membranes were blocked followed by incubation with the indicated primary antibody overnight at 4 °C and by fluorescence dye-linked secondary antibodies at room temperature for 1 hour. Immunoblots were digitally imaged using a ChemiDoc MP (Bio-Rad). Exposure time was optimized to avoid saturation. Bands were automatically defined and intensities were determined using the built-in band detection tool of the Image Lab v6.0 software (Bio-Rad).

Immunoprecipitation

Subject and control fibroblasts were harvested in ice-cold lysis buffer (50 mM HEPES, pH 7.5; 150 mM NaCl; 1%

NP-40; 0.5% sodium deoxycholate; 0.05% SDS; 1 mM EDTA) supplemented with Mini Protease Inhibitor and PhosSTOP (Roche) and lysed for 10 minutes on ice. Cell debris was removed by centrifugation for 10 minutes. Mouse anti-SF3B1 or normal mouse IgG was coupled to Dynabeads Protein G (ThermoFisher Scientific) in PBST for 10 minutes at room temperature under rotating conditions. Cleared lysates were added to the beads and incubated over night at 4 °C on a rotator. Beads were washed 5 times with ice-cold lysis buffer. After final washing, beads were resuspended in sample buffer and subjected to sodium dodecyl-sulfate polyacrylamide gel electrophoresis and immunoblotting.

Data analysis and statistics

Quantitative data are presented by GraphPad Prism 8 software (Instat, GraphPad Software) as the mean ± standard deviation (SD). For quantification one-way analysis of variance followed by a Dunnett post hoc test for multiple comparisons was performed for all experiments. A P value $< .05$ was considered statistically significant (*$P \leq .05$; **$P \leq .01$; ***$P \leq .001$; ****$P \leq .0001$).

Results

De novo *PHF5A* variants in individuals with congenital malformations and developmental delay

Through international collaborations, the Deciphering Developmental Disorders Study,[49,50] and GeneMatcher,[51] we recruited 9 subjects with variable congenital malformations and developmental delay who all carried a de novo heterozygous variant in *PHF5A* (Table 1 and Figure 1A). Detailed clinical descriptions can be found in the Supplemental Material. In summary, the 9 subjects had variable phenotypes. Prenatally, 6 subjects had intrauterine growth retardation. All subjects had motor and speech delay and developmental delay. Five subjects developed intellectual disability later in life, with subjects 3, 5, 6, and 7 developing mild intellectual disability and subject 9 developing severe intellectual disability and microcephaly. Congenital abnormalities comprised hypospadias in 3 of 4 male subjects and heart defects, inguinal hernia, and sacral dimple in 3 subjects each. Six of the 9 subjects had short stature (Table 1). Craniofacial dysmorphism is variable in the 9 subjects. The most consistent features are a high forehead and preauricular skin tag(s) in 5 subjects each (Table 1 and Figure 1B). Quantitative phenotypic analysis with PhenoScore detected a distinct phenotypic entity when investigating all 9 subjects included in this study (area under the curve 0.89, $P < .001$), whereas the analysis did not reveal significant results (area under the curve 0.62, $P = .17$) when excluding the 3 subjects with a *PHF5A* missense variant, indicating a possible power issue there.

Subject 1 carried a start-loss variant c.2T>C / p.(Met1?). Five individuals had a likely LOF variant, including a 2-bp deletion c.69_70del / p.(Cys23*) in the unrelated subjects 2 and 3, a nonsense variant c.70G>T / p.(Glu24*) in subject 4, a splice site variant c.243+1G>A in subject 5, and a 1-bp duplication c.276dup / p.(Ser93Glufs*3) in subject 6. The remaining 3 subjects carried a missense variant, with c.11A>G / p.(His4Arg) in subject 7, c.44C>T / p.(Ala15Val) in subject 8, and c.185G>A / p.(Gly62Glu) in subject 9 (Table 1). The variants c.243+1G>A in subject 5 and c.44C>T / p.(Ala15Val) in subject 8 have been previously reported in the Deciphering Developmental Disorders Study (https://www.deciphergenomics.org/ddd/research-variant/883d75e620de3e6570b2d31adec4795f/overview).[50] The *PHF5A* variants in subjects 1 to 4 and 9 were validated in fibroblast-derived DNA by Sanger sequencing (Supplemental Figure 1). All *PHF5A* variants were absent in the gnomAD database (v2.1.1 and 3.1.2). The variant c.2T>C affects the translation initiation codon of *PHF5A*. By analyzing published ribosome sequencing data of *PHF5A* (ENST00000216252.4/NM_032758.4) from human fibroblasts, no other translational initiation sites downstream of the first ATG are used for protein biosynthesis (Supplemental Figure 2).[54,55] The variant c.243+1G>A affects the highly conserved splice donor site of intron 3. In silico splice site predictions using the algorithms SpliceSiteFinder-like, MaxEntScan, NNSPLICE, and GeneSplicer predicted disruption of the canonical splice site by the intronic variant.[56-60] Because we included subject 5 with the non-coding *PHF5A* variant during the revision of the manuscript, we could not analyze the effect of the intronic variant on pre-mRNA splicing experimentally. The meta pathogenicity predictor Combined Annotation-Dependent Depletion predicted a possible deleterious effect for the *PHF5A* missense variants p.(His4Arg), p.(Ala15Val), and p.(Gly62Glu) (Supplemental Table 2). The 3 changes affect codons that are highly intolerant to genetic variation as predicted by MetaDome[61] (Supplemental Table 2). Although the 2 residues histidine 4 and glycine 62 are highly evolutionary conserved, alanine 15 shows only partial conservation (Supplemental Figure 3).[5]

We explored the structural impact of the 3 missense variants by using a three-dimensional crystallographic structure of PHF5A (Figure 1C and Supplemental Figure 4).[10] p.(Ala15Val) and p.(Gly62Glu) were predicted to affect adjacent zinc finger motifs (Figure 1C). Replacement of Ala15 (Figure 1C, panel i) by valine is predicted to cause steric clashes among side chains of valine 15 and arginine 44 (Figure 1C, panel ii). Arg44, in turn, is next to Cys46, which is 1 of 4 cysteines of zinc finger motif 1 (Cys11, Cys46, Cys49, and Cys85). Changed positioning of Arg44 might have an effect on the zinc finger motif 1 (Figure 1C, panel ii). Gly62 is in close proximity to zinc finger motif 2 (Cys23, Cys26, Cys58, and Cys61) (Figure 1C, panel iii). Change of Gly62 by glutamate was predicted to lead to steric clashes with side chains of tyrosine 54, possibly altering the positioning of zinc finger motif

Table 1 Clinical features of subjects with heterozygous *PHF5A* variants

Subject Characteristics										Total
ID	1	2	3	4	5	6	7	8	9	
Sex	Female	Male	Male	Male	Female	Female	Female	Male	Female	
Age at last examination	5 years 10 months	21 years	7 years 11 months	5 years 8 months	11 years	19 years	17 years	24 years	6 years 6 months	
PHF5A variant										
Genomic DNA (GRCh38; NC_000022.11)	g.41468652A>G	g.41468134_41468135del	g.41468134_41468135del	g.41468130C>A	g.41467447C>T	g.41460458dup	g.41468643T>C	g.41468610G>A	g.41467506C>T	
Coding DNA (NM_032758.4)	c.2T>C	c.69_70del	c.69_70del	c.70G>T	c.243+1G>A	c.276dup	c.11A>G	c.44C>T	c.185G>A	
Protein (NP_116147.1)	p.(Met1?)	p.(Cys23*)	p.(Cys23*)	p.(Glu24*)	p.?	p.(Ser93Glufs*3)	p.(His4Arg)	p.(Ala15Val)	p.(Gly62Glu)	
Inheritance	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo	
Prenatal and neonatal history										
Intrauterine growth retardation (HP:0001511)	−	+	+	−	+	+	+	−	+	6/9 (67%)
Caesarian section (HP:0011410)	+	+	+	+	−	+	+	−	−	6/9 (67%)
Premature birth (HP:0001622)	+	+	+	−	−	+	−	−	+	5/9 (56%)
Respiratory failure requiring assisted ventilation (HP:0004887)	−	+	+	−	−	+	−	−	−	3/9 (33%)
Congenital anomalies										
Hypospadias (HP:0000047)	NA	+	+	−	NA	NA	NA	+	NA	3/4 (75%)
Abnormal aortic arch morphology (HP:0012303)	−	−	−	+	−	−	−	−	−	1/9 (11%)
Atrial septal defect (HP:0001631)	−	−	−	−	−	−	−	+	+	2/9 (22%)
Ventricular septal defect (HP:0001629)	−	−	−	+	−	−	−	−	−	1/9 (11%)

(continued)

Table 1 Continued

Subject Characteristics										Total
Inguinal hernia (HP:0000023)	−	+	+	−	−	−	−	+	−	3/9 (33%)
Sacral dimple (HP:0000960)	+	+	−	−	−	−	−	+	−	3/9 (33%)
Growth abnormalities										
Small for gestational age (HP:0001518)	−	+	ND	−	+	ND	+	−	+	4/7 (57%)
Microcephaly (HP:0000252)	−	+	−	−	−	−	−	−	+	2/9 (22%)
Macrocephaly (HP:0000256)	+	−	−	−	−	−	−	−	−	1/9 (11%)
Short stature (HP:0004322)	−	+	−	−	+	+	+	+	+	6/9 (67%)
Development										
Motor delay (HP:0001270)	+	+	+	+	+	+	+	+	+	9/9 (100%)
Speech delay (HP:0000750)	+	+	+	+	+	+	+	+	+	9/9 (100%)
Intellectual disability (HP:0001249)	−	−	+	−	+	+	+	−	+	5/9 (56%)
Intellectual disability, mild (HP:0001256)	−	−	+	−	+	+	+	−	−	4/9 (44%)
Neurological abnormalities										
Hypotonia (HP:0001252)	−	−	+	−	+	−	−	+	+	4/9 (44%)
Seizure (HP:0001250)	−	−	+	−	−	−	−	−	−	1/9 (11%)
Behavioral abnormalities										
Autistic behavior (HP:0000729)	−	−	+	−	−	−	−	−	−	1/9 (11%)
Sleep disturbance (HP:0002360)	−	−	+	−	−	−	−	−	−	1/9 (11%)

(continued)

Table 1 Continued

Subject Characteristics										Total
Skeletal abnormalities										
Talipes equinovarus (HP:0001762)	−	−	−	−	+	−	−	−	−	1/9 (11%)
Postaxial polydactyly (HP:0100259)	−	−	−	+	−	−	−	−	−	1/9 (11%)
Long fingers (HP:0100807)	−	−	−	−	+	+	+	−	−	3/9 (33%)
Broad toes (HP:0001837)	−	+	−	−	+	−	−	+	−	3/9 (33%)
Short distal phalanx of hallux (HP:0010103)	−	−	−	−	−	−	−	+	−	1/9 (11%)
Other abnormalities										
Small nails (HP:0001792) and/or nail dysplasia (HP:0002164)	−	−	−	−	+	−	−	+	−	2/9 (22%)
Dysmorphic features										
High forehead (HP:0000348)	+	+	+	+	−	−	+	−	−	5/9 (55%)
Preauricular skin tag (HP:0000384)	−	+	+	+	+	+	−	−	−	5/9 (55%)
Periorbital fullness (HP:0000629)	+	−	−	+	+	−	+	−	−	4/9 (44%)
Deeply set eyes (HP:0000490)	−	−	+	+	−	−	+	−	−	3/9 (33%)
Hypertelorism (HP:0000316)	+	−	−	−	+	+	−	−	−	3/9 (33%)
Downslanted palpebral fissures (HP:0000494)	−	−	+	−	−	−	−	+	−	2/9 (22%)
Low-set ears (HP:0000369)	−	−	−	+	+	+	−	−	+	4/9 (44%)
Retrognathia (HP:0000278)	−	−	−	−	−	+	−	−	+	2/9 (22%)

+, present; −, absent; *NA*, not applicable; *ND*, not documented.

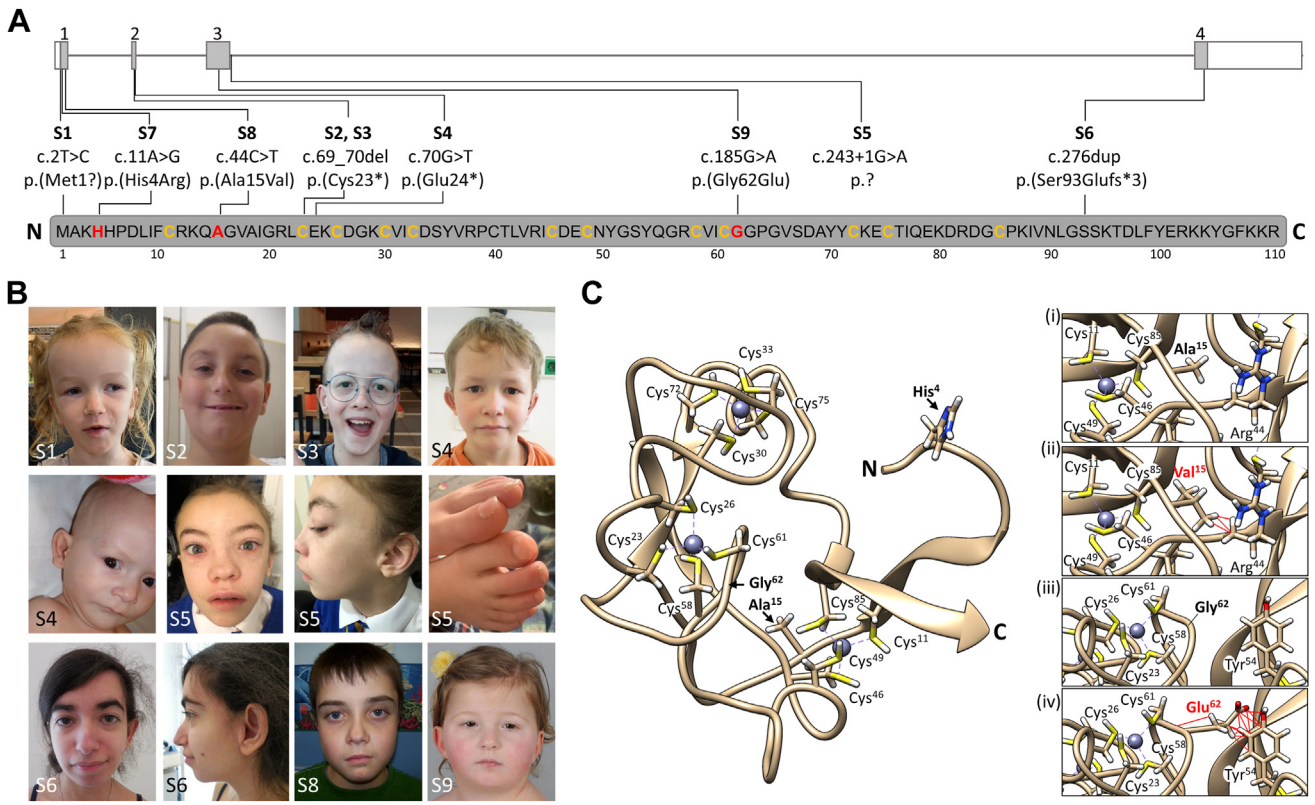


Figure 1 ***PHF5A* variants, photographs of 8 affected individuals, and molecular modelling of PHF5A.** A. Location of the de novo *PHF5A* variants in the exon-intron-structure (NM_032758.4) and the amino acid sequence of PHF5A (NP_116147.1). Numbering of exons is given. Untranslated region in exons is indicated in white and coding region in gray. The c.69_70del variant was identified in 2 unrelated subjects. Cysteine residues are highlighted in yellow in the amino acid sequence and residues affected by missense variants in red. PHF5A contains 3 zinc-finger motifs, each consisting of 4 cysteine residues ((1) Cys^{11}, Cys^{46}, Cys^{49}, and Cys^{85}; (2) Cys^{23}, Cys^{26}, Cys^{58}, and Cys^{61}; (3) Cys^{30}, Cys^{33}, Cys^{72}, and Cys^{75}). B. Photographs of 8 subjects with a de novo *PHF5A* variant. From left to right: Upper row, subject 1 with c.2T>C / p.(Met1?) at age 6 years and 6 months, subject 2 with c.69_70del / p.(Cys23*) at age 10 years, subject 3 with c.69_70del / p.(Cys23*) at age 11 years and subject 4 with c.70G>T / p.(Glu24*) at age 6 years and 9 months; middle row, subject 4 with c.70G>T / p.(Glu24*) at 5 months and photographs of face, ear, and toes of subject 5 with c.243+1G>A / p.? at 11 years. Subject 5 has dry flaky skin over her eyelids and nose and nail dysplasia; lower row: subject 6 with c.276dup / p.(Ser93Glufs*3) at age 19 years, subject 8 with c.44C>T / p.(Ala15Val) at age 10 years, and subject 9 with c.185G>A / p.(Gly62Glu) at age 4 years and 6 months. Dysmorphic features include a high forehead, down slanted palpebral fissures, and low-set ears. Preauricular skin tags are shown in subjects 4 to 6. C. Structural impact of *PHF5A* missense variants. Model of PHF5A (PDB ID: 5SYB) published by Teng et al. (2017).[10] Molecular graphics were developed using UCSF Chimera.[52] PHF5A backbone is shown as beige ribbon and His^4, Ala^{15}, Gly^{62}, and cysteines coordinating zinc in the zinc finger motifs as sticks. Zinc atoms are depicted as gray spheres. Close-ups for Ala^{15} (i) and Gly^{62} (iii) and surrounding amino acids and adjacent zinc finger motif are shown on the right. Amino acid substitutions with Val at position 15 (red) (ii) and Glu at position 62 (red) (iv) was simulated using Chimera's rotamers tool.[53] Steric clashes among amino acid side chains with a distance ≤ 0.6 Å are represented by red lines. C, C terminus; N, N terminus; S, subject.

2 (Figure 1C, panel iv). No steric effect could be predicted for the p.(His4Arg) substitution through structural modelling (Supplemental Figure 4). *PHF5A* is intolerant to LOF variants as the loss of function observed/expected upper bound fraction (LOEUF) is 0.46 (gnomAD v2.1.1). This constraint metric correlates with the mean LOEUF of 0.488 for 389 orthologous genes that causes embryonic lethality after heterozygous deletion in mouse and the mean LOEUF of 0.636 for 678 genes required for human cell viability.[62] The Z score of 2.65 for missense variants is slightly increased and there are only 5 observed versus 65.4 expected missense variants in *PHF5A* (~8%) in gnomAD v2.1.1.[63] Together, based on the absence of the identified *PHF5A* variants in population databases, 3 LOF and 2 potential LOF variants, all de novo, in 6 subjects with overlapping clinical features, we were convinced that these heterozygous *PHF5A* variants underlie the phenotype in the affected individuals.

PHF5A transcripts expressed from the LOF allele are stable, and *PHF5A* mRNA levels are similar in subject- and control-derived fibroblasts

To test the possibility that some of the *PHF5A* LOF variants lead to nonsense-mediated mRNA decay (NMD) of *PHF5A*

variant transcripts, we applied the in silico program NMDEscPredictor.[64] This tool predicted the *PHF5A* variant c.69_70del to be subject to degradation by NMD, whereas the variant c.276dup was predicted to escape NMD. To experimentally test the predictions, we first performed qualitative RT-PCR followed by Sanger sequencing using fibroblast-derived cDNA of subjects 1 to 4 and 9 and identified both wild-type and variant *PHF5A* transcripts in cells with a heterozygous *PHF5A* LOF or start-loss variant, as well as in cells with the c.185G>A / p.(Gly62Glu) variant (Figure 2A). The ratio of wild-type to variant *PHF5A* transcripts was about 1:1 in cells of subjects 1, 2, 4, and 9, whereas the sequence traces of variant *PHF5A* mRNAs in subject 3 cells were approximately twice that of wild-type mRNAs (Figure 2A). The data suggest that *PHF5A* transcripts bearing a premature stop codon are stable in subject-derived fibroblasts and not subject to NMD.

We next determined relative *PHF5A* mRNA levels in fibroblasts of subjects 1 to 4 and 9 by RT-qPCR. The *PHF5A* mRNA amount was similar in fibroblasts of subjects 1, 2, 4, and 9 and controls (Figure 2B and C). In contrast, the *PHF5A* transcript level was significantly increased by 1.5- to 1.7-fold in fibroblasts of subject 3 carrying the *PHF5A* variant c.69_70del compared with control cells (Figure 2D). We also performed transcriptome sequencing in fibroblasts of subjects 1, 2, 4, and 9 and obtained similar normalized *PHF5A* RNA read counts in the 4 affected individuals compared with the control cohort (42 fibroblast cell lines) (Figure 2E and F). Together, these data indicate the presence and stability of *PHF5A* variant transcripts in fibroblasts of subjects carrying a heterozygous *PHF5A* LOF allele that likely contributed to normal *PHF5A* mRNA levels. In fibroblasts of subject 3 with the *PHF5A* c.69_70del variant, an increased *PHF5A* mRNA amount was identified that was not detected in fibroblasts of subject 2 carrying the same *PHF5A* variant. Although these data could be explained by an interindividual variability of *PHF5A* expression in fibroblast cells, an individual compensatory effect at the transcriptional level in subject 3-derived cells may also be a possible explanation.

We evaluated transcriptome sequencing data from fibroblasts of subjects with a *PHF5A* LOF variant (subjects 2 and 4) and the start-loss variant (subject 1) for global changes in the number of alternative splicing events, such as alternative 5′ and 3′ splice site selection, alternative first or

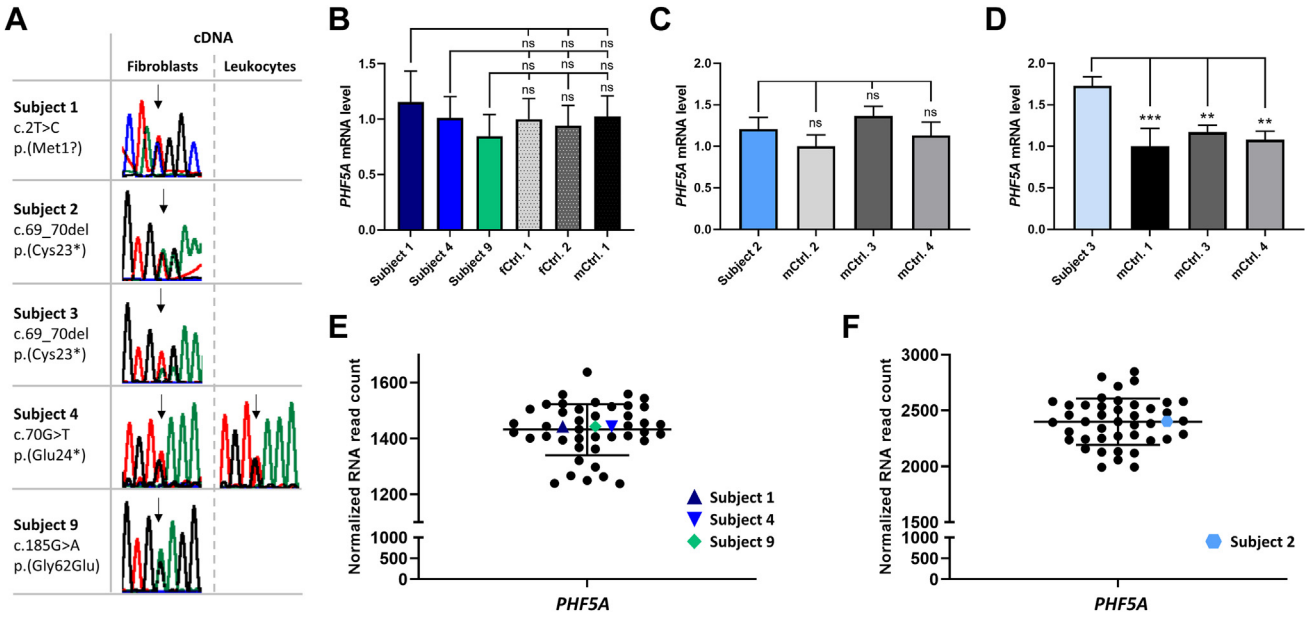
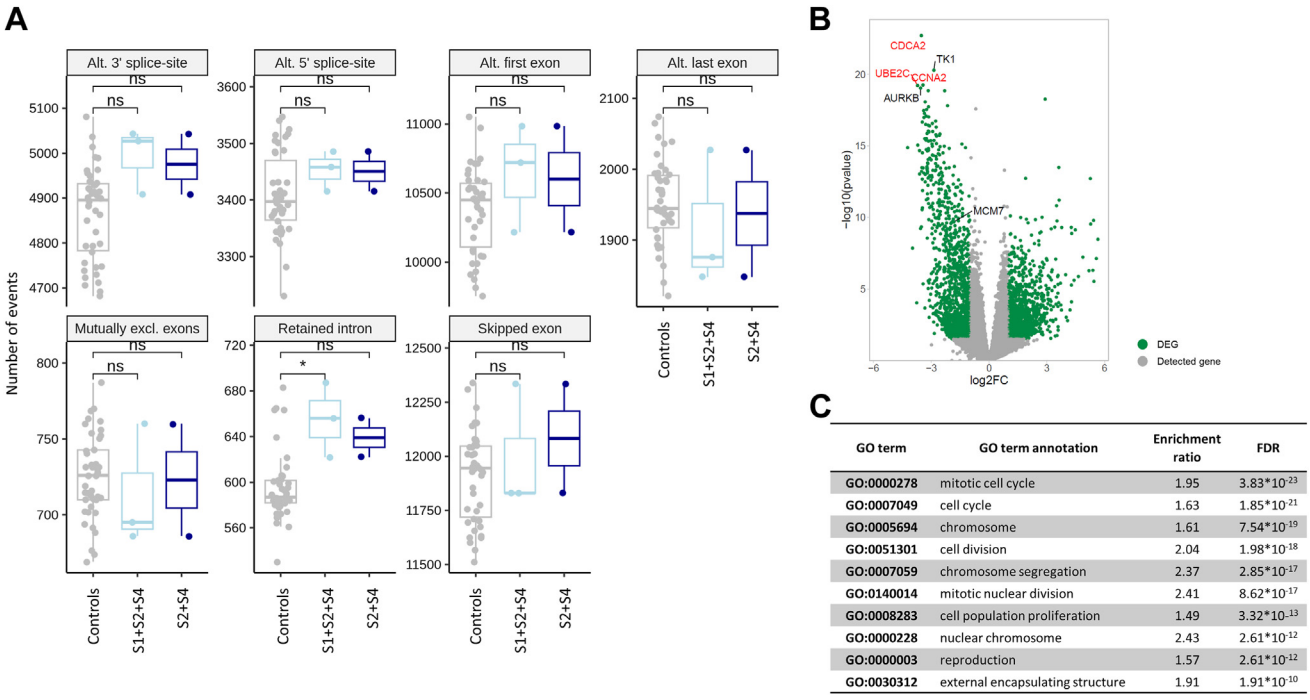


Figure 2 Qualitative and quantitative *PHF5A* transcript analysis in subject-derived fibroblasts and/or leukocytes. A. Partial sequence electropherograms show the presence of *PHF5A* variant-bearing transcripts in fibroblast-derived cDNA of subjects 1-4 and 9 (middle column) and in leukocyte-derived cDNA of subject 4. The ratio of wild-type versus variant *PHF5A* transcripts was about 1:1 in subjects 1, 2, 4, and 9. In fibroblast-derived cDNA of subject 3, the Sanger sequence profile showed a reduced signal for the aberrant transcripts with the 2-bp deletion superimposed on the wild-type sequence with an estimated ratio of wild-type versus variant *PHF5A* transcripts of ~2:1. Arrows point to the position/start of the variant. B-D. Relative quantification of *PHF5A* mRNA levels by real-time quantitative PCR using fibroblast-derived cDNA of subjects 1-4 and 9 and indicated controls normalized to *GAPDH* mRNA levels. Bars represent the mean ± SD of 4 (B), 2 (C), or 3 (D) independent experiments, each performed in triplicate. One-way ANOVA followed by Dunnett post hoc test was performed separately between individual subjects and controls. $^{**}P \leq .01$; $^{***}P \leq .001$. E and F. Relative *PHF5A* mRNA expression in fibroblasts of subjects 1, 4, and 9 (E) and of subject 2 (F) by transcriptome sequencing. The plots show normalized *PHF5A* RNA read counts in the transcriptome data set of 42 fibroblast control and 3 (E) or 1 (F) subject cell line(s). RNA of subject 2 fibroblasts was sequenced on another sequencing run as RNA of subject 1, 4, and 9 fibroblasts. Therefore, subject-2-derived RNA read counts were normalized separately to controls. The mean ± SD normalized read counts of 45 (E) or 43 (F) samples is shown. cDNA, complementary DNA; fCtrl. 1 and 2, female control fibroblasts; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; mCtrl. 1-4, male control fibroblasts; mRNA, messenger RNA; ns, not significant; PCR, polymerase chain reaction.

last exons, mutually exclusive exons, retained introns, and skipping of exons. When comparing the group of subjects with clear LOF variants (subjects 2 and 4) against the group of controls (42 fibroblast cell lines), we did not find significantly disturbed alternative splicing integrity in fibroblasts of the 2 individuals with *PHF5A* LOF variants compared with control fibroblasts (Figure 3A). When adding data of subject 1 with the *PHF5A* start-loss variant, a potential LOF variant, to the subject cohort, the number of retained introns was slightly but significantly increased in the subject group compared with the control group (Figure 3A). Nevertheless, the number of retained introns in subjects 1, 2, and 4 was comparable to the number of retained introns in outliers of the control group (Figure 3A). We next looked into alternative splicing events of individual genes in subjects compared with controls. When we compared data from subjects 2 and 4 with *PHF5A* LOF variants with the control group, we identified 15 significantly different splicing events in subjects (Supplemental Figure 5 and Supplemental Table 3). None of them was a retained intron event. Instead, an alternative first exon was used in 12 events, a cryptic cassette exon in 2, and an alternative 5′ splice site in 1 event (Supplemental Table 3). Inclusion of subject 1 in the subject group resulted in the detection of 14 significantly different splicing events compared with the control group, of which 12 were alternative first exon usage, 1 alternative last exon usage, and 1 usage of an alternative 5′ splice (Supplemental Figure 6 and Supplemental Table 4). Ten significantly different splicing events representing alternative first exon and alternative 5′ splice site usages were shared in the analyses of 2 (subjects 2 and 4) or 3 individuals (subjects 1, 2, and 4) in the subject group compared with the control group (Supplemental Table 4). Using the online tool Enrichr,[65] we investigated whether gene products of the identified genes with significant alternative splicing events share biological pathways. Significantly enriched Reactome pathways of genes identified in subjects 2 and 4 were RHO GTPase-, tRNA processing-, and cell-cycle-related pathways (Supplemental Figure 7A). For genes with significant alternative splicing events identified in subjects 1, 2, and 4, we found tRNA processing- and cell-cycle-related pathways enriched



GO term	GO term annotation	Enrichment ratio	FDR
GO:0000278	mitotic cell cycle	1.95	$3.83*10^{-23}$
GO:0007049	cell cycle	1.63	$1.85*10^{-21}$
GO:0005694	chromosome	1.61	$7.54*10^{-19}$
GO:0051301	cell division	2.04	$1.98*10^{-18}$
GO:0007059	chromosome segregation	2.37	$2.85*10^{-17}$
GO:0140014	mitotic nuclear division	2.41	$8.62*10^{-17}$
GO:0008283	cell population proliferation	1.49	$3.32*10^{-13}$
GO:0000228	nuclear chromosome	2.43	$2.61*10^{-12}$
GO:0000003	reproduction	1.57	$2.61*10^{-12}$
GO:0030312	external encapsulating structure	1.91	$1.91*10^{-10}$

Figure 3 Number of splice events and differentially expressed genes in transcriptome sequencing data of subject fibroblasts. A. Splice events in transcriptome data were retrieved from human genome assembly GRCh38, GENCODE version 39. The proportion-spliced-in (PSI) was determined using SUPPA2. Events with a PSI ≥ 20% were counted. The number of splice events is shown by groups comparing either the group of subjects with *PHF5A* LOF variants (subjects 2 and 4; dark blue box plots) alone or together with subject 1 carrying a start-loss variant (subjects 1, 2, and 4; light blue boxplots) against the group of 42 controls (grey box plots). Differences between subject groups and the control group were tested using Wilcoxon rank-sum tests separately for each event type: $*P \leq .05$. B. Volcano plot of differentially expressed genes. Differentially expressed genes between subjects 2 and 4 with *PHF5A* LOF variants and the set of 42 controls were determined using DESeq2. A total of 2622 genes showed significant differential expression between subject and control group (green dots), including 1248 upregulated and 1374 downregulated genes. *CDCA2*, *TK1*, *CCNA2*, *UBE2C*, and *AURKB* are the top 5 downregulated genes in the subject group. Data point for *MCM7* is indicated as this gene was among the 10 genes with alternative promoter usage and was downregulated in subject cells (Supplemental Tables 3 and 4). C. Gene ontology (GO) term analysis of genes significantly differentially expressed between subject and control group. Top 10 most significantly enriched GO terms are listed by false discovery rate. Alt, alternative; DEG, differentially expressed gene; excl., exclusive; FDR, false discovery rate; GO, gene ontology; ns, not significant; S, subject.

(Supplemental Figure 7B). Shared genes in the 2 analyses (subjects 2 and 4 or subjects 1, 2, and 4 versus control group) were *TRMU* und *CPSF4* involved in tRNA processing and *FBXO5* and *MCM7* implicated in cell-cycle regulation (Supplemental Figure 7).

We next investigated differentially expressed genes (DEGs) in the transcriptome sequencing data of subjects 1, 2, and 4. Comparison of transcriptome data of subjects 2 and 4 with *PHF5A* LOF variants with the set of 42 controls resulted in 2622 DEGs (1248 upregulated and 1374 downregulated genes) (Figure 3B). Comparison of transcriptome data from fibroblasts of the 3 subjects 1, 2, and 4 with the set of 42 controls led to 3374 DEGs (1846 upregulated and 1528 downregulated genes) (Supplemental Figure 8A). In the 2 analyses (subjects 2 and 4 or subjects 1, 2, and 4 vs control group), *CCNA2*, *CDCA2*, and *UBE2C* were shared genes among the top 5 downregulated genes (Figure 3B and Supplemental Figure 8A). Of the genes for which significantly different splicing events were identified, *MCM7* and *FBXO5* were additionally significantly downregulated to approximately 27% and 20%, respectively, in fibroblast cells of subjects 1, 2, and 4 compared with controls (Supplemental Tables 3 and 4). We analyzed the biological role of all detected DEGs. Significantly enriched gene ontology (GO) terms, independent of inclusion of subject 1 in the subject group or not, were found to be associated with mitosis and cell cycle (Figure 3C and Supplemental Figure 8B).

Formation of the U2 SF3B splicing complex is not altered in fibroblasts of subjects with a *PHF5A* pathogenic variant

In fibroblasts of subjects with heterozygous *PHF5A* LOF variants, both wild-type and variant *PHF5A* transcripts likely gave rise to normal *PHF5A* mRNA levels. However, we expect that *PHF5A* mRNAs bearing an early premature stop codon or start codon loss do not lead to production of the full-length PHF5A protein. Instead, the 4 heterozygous *PHF5A* LOF variants are expected to be associated with haploinsufficiency resulting in a 50% reduction of PHF5A protein levels. To study this, we determined PHF5A levels in whole-cell lysates of fibroblasts of subjects and controls in immunoblotting by applying an anti-PHF5A antibody which has been generated using the full-length human PHF5A protein. We detected PHF5A with a molecular mass of approximately 15 kDa in cells of subjects 1 to 4 and 9 and 6 controls (Figure 4A and Supplemental Figure 9A). In the 6 control cell lines, we observed interindividual variability of PHF5 levels (Figure 4A and Supplemental Figure 9A). No C-terminally truncated versions of PHF5A were detected in subject 2 to 4 cells with a heterozygous LOF variant upon longer exposure of immunoblots (Supplemental Figure 10). Quantification of PHF5A values revealed similar PHF5A levels in fibroblasts of subjects 1, 3, 4, and 9 and controls, whereas the PHF5A level in fibroblasts of subject 2 with the p.(Cys23*) variant was significantly reduced to 43% to 54% compared with controls (Figure 4B and Supplemental Figure 9B). These data demonstrate that wild-type PHF5A (molecular mass of ~15 kDa) levels were variable but almost normal (between 60% and 104%) in fibroblasts of subjects with a heterozygous *PHF5A* variant, including LOF variants, except of subject 2. Similar PHF5A levels in control and subject 9 cells with the heterozygous missense variant p.(Gly62Glu) could be expected. However, these results were unexpected in cells with (1) the heterozygous *PHF5A* start-loss variant (subject 1) because there is no second methionine in the PHF5A amino acid sequence (Figure 1A) and (2) the heterozygous LOF variants c.69_70del / p.(Cys23*) (subject 3) and c.70G>T / p.(Glu24*) (subject 4) because *PHF5A* mRNAs with a premature stop codon cannot lead to the production of full-length PHF5A. Only subject 2 fibroblasts with the stop-gain p.(Cys23*) variant had the expected about 50% reduced PHF5A level that is suggestive of normal PHF5A protein production from wild-type *PHF5A* mRNAs expressed from the single intact gene copy. In subject 3 and 4 cells with a heterozygous *PHF5A* LOF allele, the absence of a 50% reduction of wild-type PHF5A levels is suggestive of an upregulated translation or decreased degradation of wild-type *PHF5A* transcripts. This could represent a compensatory mechanism to bring the PHF5A protein level to an almost normal cellular amount. Although this cellular response could be present in fibroblasts of subject 3 with the p.(Cys23*) variant, it was not observed in fibroblasts of subject 2 carrying the same *PHF5A* variant.

To analyze the possibility of an intrinsic instability of PHF5A p.(Ala15Val) and p.(Gly62Glu) variant proteins, we used pIRES2-EGFP-PHF5A constructs enabling simultaneous translation of GFP and N- or C-terminally FLAG-tagged PHF5A wild type and variants from a single transcript in transiently transfected HEK293T cells. The p.(His4Arg) variant was not included in this analysis because subject 6 with this *PHF5A* variant was included during the revision process. The GFP amount was similar in HEK293T cells expressing FLAG-tagged wild-type and PHF5A variants (Supplemental Figure 11A), indicating the same transfection efficiency for all *PHF5A* expression constructs. We did not detect any significant difference in the amount of ectopically expressed PHF5A wild type and both variants, independent of the location of the FLAG-tag (Supplemental Figure 11B).

PHF5A is part of the SF3B splicing complex consisting of SF3B1, SF3B2, SF3B3, SF3B4, SF3B5, and SF3B6.[5-7] In a first step to analyze the integrity of the SF3B complex in subject-derived fibroblasts, we determined levels of SF3B1, SF3B2, SF3B3, and SF3B6 in fibroblasts of subjects 1 to 4 and 9. Immunoblotting followed by quantitative analysis revealed that the levels of all 4 SF3B components in fibroblasts of subjects 1, 3, 4, and 9 were comparable to those of control cells (Figure 4C-F and Supplemental Figure 9C-F). In cells of subject 2 with the p.(Cys23*) variant, which showed reduced PHF5A protein levels as

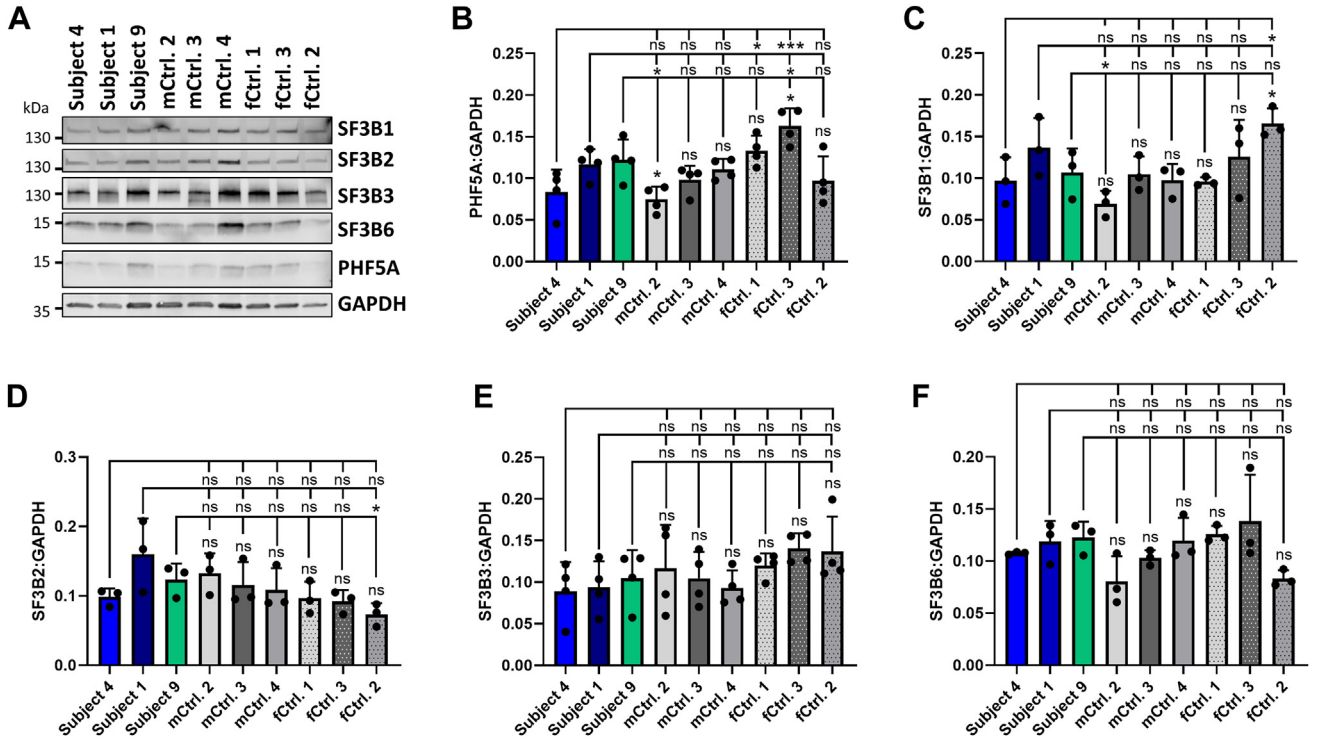


Figure 4 Protein levels of PHF5A, SF3B1-3, and SF3B6 in fibroblasts of subjects 1, 4, and 9. A. Representative immunoblots of whole-cell lysates obtained from subject and control fibroblast cultures from 3 to 4 different passages. Levels of PHF5A, SF3B1, SF3B2, SF3B3, and SF3B6 were monitored with the indicated antibodies. An anti-GAPDH antibody was used to demonstrate equal loading. B-F. Band intensities of fluorescence signals were quantified using the ChemiDoc imaging system. PHF5A, SF3B1-3, and SF3B6 protein levels were normalized to GAPDH. B, E. The mean ± SD of 4 independent experiments is shown. C, D, and F. The mean ± SD of 3 independent experiments is shown. One-way ANOVA followed by Dunnett post hoc test was performed separately between individual subjects and controls. *$P \leq .05$; ***$P \leq .001$. fCtrl. 1-3, female control fibroblasts; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; mCtrl. 2-4, male control fibroblasts; ns, not significant.

discussed above (Supplemental Figure 9B), varying levels of SF3B complex proteins were identified. Subject 2 fibroblasts had significantly reduced levels of SF3B2, SF3B3, and SF3B6 compared with 2 or all 3 control cell lines, whereas the SF3B1 level was unaffected (Supplemental Figure 9C-F). These data show the different capabilities of subject-derived fibroblasts to maintain normal levels of SF3B complex proteins. In subject-derived fibroblasts with normal PHF5A levels, amounts of other SF3B complex components were similar to control cells. Remarkably, low levels of PHF5A were accompanied by diminished levels of SF3B2, SF3B3, and SF3B6 in subject 2-derived fibroblasts.

We next analyzed formation of the SF3B complex in 2 selected subject-derived fibroblast cell lines. We used fibroblasts of subject 4 with the *PHF5A* variant p.(Glu24*) and of subject 9 with the missense change p.(Gly62Glu). p.(Gly62Glu) may have a negative impact on complex formation because of decreased interaction with SF3B1 or any of the other SF3B complex proteins. We immunoprecipitated endogenous SF3B1 and detected co-immunoprecipitated endogenous SF3B2, SF3B3, SF3B6, and PHF5A in the immunoprecipitates. All 4 SF3B complex components efficiently co-immunoprecipitated with SF3B1 in fibroblasts of subjects 4 and 9 and controls (Supplemental Figures 12A and 13A). Quantitative analysis revealed similar amounts of co-immunoprecipitated SF3B2, SF3B3, SF3B6, and PHF5A with SF3B1 in subject- and control-derived fibroblasts (Supplemental Figures 12B-E and 13B-E). Our data demonstrate that formation of the SF3B subcomplex of the U2 snRNP is not altered in fibroblasts of subjects with heterozygous *PHF5A* variants. Moreover, the *PHF5A* p.(Gly62Glu) variant does not affect assembly of the SF3B complex in subject 9 cells.

Discussion

Here, we describe 9 subjects with de novo heterozygous variants in *PHF5A*, including 4 with LOF variants, 3 with a missense variant, 1 with a splice variant, and 1 with a start-loss variant. Although there is no other available ATG codon in the *PHF5A* coding region,[54,55] usage of alternate translation initiation sites for PHF5A protein biosynthesis from mRNAs harboring the c.2T>C variant in cells from subject 1 cannot be ruled out. The splice variant likely leads to aberrantly spliced *PHF5A* mRNAs. The associated phenotype included craniofacial dysmorphism, particularly preauricular skin tag(s), developmental delay, hypospadias,

and pre- and post-natal growth abnormalities. Using quantitative next-generation phenomics, we show the phenotype is distinguishable from individuals with other neurodevelopmental disorders. After the recent report of a de novo *PHF5A* nonsense variant c.162C>A / p.(Tyr54*) in a girl with left microtia and bilateral absence of 12th ribs,[38] the association of craniofacial and/or skeletal anomalies with heterozygous *PHF5A* LOF variants is confirmed.

The effect of the 3 missense variants p.(His4Arg), p.(Ala15Val), and p.(Gly62Glu) on PHF5A function is as yet unclear. In fibroblasts with the heterozygous p.(Gly62Glu) variant, complex formation between PHF5A and other SF3B components was unaffected. The 2 missense variants p.(Ala15Val) and p.(Gly62Glu) did not seem to affect intrinsic stability of respective PHF5A variant proteins when expressed in a heterologous cell system. Thus, the *PHF5A* missense variants p.(His4Arg), p.(Ala15Val), and p.(Gly62Glu) are classified as variants of uncertain clinical significance with further evidence needed to demonstrate pathogenicity.

Analysis of transcriptome sequencing data in fibroblasts of subjects 1, 2, and 4 for global changes in splice events identified a slight increase in the number of retained introns; however, we did not observe intron retention in specific/shared genes nor a severely disturbed splicing integrity in the 3 subject cell lines. More detailed analysis, including alternative splice events for individual genes, identified alternative promoter usage in 10 genes but no other significantly different splicing events in subject 1, 2, and 4 fibroblasts compared with the control group. *FBXO5* and *MCM7* were among the 10 genes with alternative promoter use. Interestingly, both genes were also significantly downregulated in fibroblasts of subjects 1, 2, and 4, in addition to *CCNA2*, *CDCA2*, and *UBE2C*, which belong to the top 5 downregulated genes in the 3 subject fibroblasts. *FBXO5*, *MCM7, CCNA2*, *CDCA2*, and *UBE2C* encode proteins with a direct or indirect role (via chromatin organization or DNA replication) in cell-cycle regulation.[66-70] This is consistent with a general enrichment of DEGs with a function in the cell cycle or cell division in subject cells. It is as yet unclear whether the enrichment of genes involved in mitosis and the cell cycle that showed an alternative promoter usage and/or downregulated expression in the 3 subject cell lines was caused by a direct effect, that is, because of the *PHF5A* variant, or an indirect effect, that is, a compensatory mechanism (see below). This needs further investigation.

Subjects with heterozygous *PHF5A*, *SF3B2*, or *SF3B4* pathogenic variants share some craniofacial anomalies, such as down slanted palpebral fissures and preauricular skin tag(s); however, the phenotype in individuals with *PHF5A* variants seems to be significantly milder than in Nager syndrome and craniofacial microsomia. Limb malformations are consistently found in Nager syndrome, whereas they are rare in individuals with *SF3B2* and *PHF5A* variants.[13,37] Taken together, the *PHF5A*-related disorder can be added to the group of craniofacial spliceosomopathies and in particular to the subgroup related to defects in the SF3B spliceosome complex.

Haploinsufficiency has been proposed to be the underlying mechanism of *SF3B2*- and *SF3B4*-related diseases because most of the affected individuals carry heterozygous LOF variants.[29,32,36,37] The presence of heterozygous null alleles in the patients suggested nonsense-mediated mRNA decay of variant transcripts and/or production of non-functional C-terminally altered proteins. Thus, in most of the cases, about 50% of wild-type protein is expected to be produced in patient-derived cells. Qualitative and/or quantitative transcript and protein analyses in patient cells with heterozygous *SF3B2* or *SF3B4* LOF variants have not, or rarely, been performed.[37,71] In subjects with the heterozygous *SF3B4* variants c.614_615insG / p.(Asp205Glufs*281) (in exon 3 of 6 exons in total) and c.1060dup / p.(Arg354Profs*132) (in exon 5) SF3B4 transcripts and proteins have been studied in fibroblasts and chondrocytes, respectively. In fibroblasts of the patient with the c.614_615insG variant, wild-type and variant *SF3B4* mRNAs with a ratio of about 2:1 were detected (Figure 2C)[71]. Total *SF3B4* transcript levels were reduced by about 30% in patient-derived fibroblasts and chondrocytes compared with controls. On SF3B4 protein level, about 80% was detected in fibroblasts with the c.614_615insG variant compared with 1 control. In the respective immunoblots, a very faint band possibly representing an altered SF3B4 protein with a higher molecular weight suggests potential production of a minimal amount of the SF3B4 p.(Asp205Glufs*281) variant protein. However, quantification of SF3B4 levels has been performed for SF3B4 with a molecular weight corresponding to wild type (compare the 2 lanes in the immunoblot of Figure 3C in [71]) that revealed a decrease of only about 20% in patient cells. Altogether, SF3B4 mRNA and protein data demonstrate (1) inefficient NMD of *SF3B4* transcripts harboring a LOF variant and (2) enhanced stability and/or increased production of wild-type SF3B4 protein from mRNAs expressed from the normal *SF3B4* allele in patient cells. These results are in accordance with our data. We identified escape of NMD for *PHF5A* variant transcripts harboring a premature termination codon in subject-derived fibroblasts and/or leukocytes. Even *PHF5A* transcripts with a premature stop codon located in exon 2, that is, c.69_70del / p.(Cys23*) that were predicted to be degraded were stable. Wild-type and variant *PHF5A* transcripts gave rise to normal or increased total *PHF5A* mRNA levels in subject cells. On protein level, we observed similar amounts of PHF5A with the expected molecular weight of the wild type in almost all subject and control fibroblasts. Only in subject 2 cells with the *PHF5A* p.(Cys23*) variant, the PHF5A level was reduced by about 50%. These data are suggestive of a feedback mechanism in most subject-derived cells expressing both *PHF5A* wild-type transcripts and transcripts with a premature termination codon that leads to wild-type PHF5A protein levels between 60% and 100%. In turn, most of the subject-derived fibroblasts with a heterozygous LOF allele were able to maintain homeostatic

levels of the 4 SF3B complex proteins SF3B1, SF3B2, SF3B3, and SF3B6. Subject 2-derived fibroblasts with low PHF5A levels (~50%) had similarly low levels of SF3B2, SF3B3, and SF3B6. In fibroblasts with the heterozygous *PHF5A* variant p.(Glu24*), SF3B complex formation was unaffected. Collectively, these data show the different capabilities of subject-derived fibroblasts with heterozygous *PHF5A* null alleles to maintain sufficient levels of wild-type PHF5A and other SF3B complex proteins for effective pre-mRNA splicing.

These feedback mechanisms seem to be crucial to the homeostatic maintenance of core spliceosomal components and underscore the essential role of PHF5A in splicing that cannot be taken over by other splicing factors. For several genes encoding splicing factors (negative) autoregulation to control their own expression has been demonstrated. Multiple layers of control do exist: (1) alternative splicing of pre-mRNAs expressed from the splicing factor gene leads to production of transcript variants containing premature termination codons that elicits NMD.[72-77] This process is important to buffer the concentration of splicing factors and to reduce cell-cell heterogeneity in splicing factor levels,[78] (2) splicing factor transcripts are subject to silencing by microRNAs,[79,80] and (3) overexpression of splicing factors decreases translational efficiency of their own mRNAs.[79] Interestingly, heterozygous non-coding variants in *SNRPB* associated with cerebrocostomandibular syndrome cause increased inclusion of an alternative exon containing a premature termination codon in *SNRPB* mRNAs. Preferential inclusion of this exon triggers NMD and leads to significantly reduced total *SNRPB* mRNA levels in patients compared with controls that is consistent with hypomorphic *SNRPB* alleles underlying this particular craniofacial spliceosomopathy.[14] In summary, negative autoregulation is important to control and maintain homeostatic levels of spliceosomal proteins, especially in a spatiotemporally specific manner during embryonic development.

The craniofacial syndromes associated with pathogenic variants in splicing factor genes are thought to be caused by decreased levels of splicing factors that reduced splicing efficiency in cranial neural crest cells, subsequently leading to neural crest cell loss.[11,13] Because this cell type forms most structures of the face, depletion of neural crest progenitors results in craniofacial abnormalities in various animal models.[81-84] However, splicing factor autoregulation may allow for multiple feedback mechanisms in cells harboring heterozygous LOF variants in genes encoding spliceosomal components, such as *SF3B4* and *PHF5A*. For example, transcripts with a premature stop codon could be retained in the nucleus and thereby escape NMD. In addition, translational efficiency of wild-type transcripts expressed from the normal allele could be increased. We put forward the hypothesis that subject-derived fibroblasts with heterozygous *PHF5A* LOF variants upregulate expression of wild-type transcripts from the normal *PHF5A* allele and/or enhance translation of wild-type transcripts to produce almost normal PHF5A levels. Translational efficiency can be modulated by shifting the target mRNA toward polyribosomes and monoribosomes to increase and decrease translation, respectively, as shown for the splicing factor SFRS1.[79] The observed compensatory mechanisms to maintain homeostatic levels of splicing factors in certain types of patient cells ([71] and this work) question haploinsufficiency as proposed pathomechanism underlying certain craniofacial spliceosomopathies.[29,37] Instead, disturbed or inefficient autoregulation of mutated spliceosomal genes by post-transcriptional and/or translational mechanisms may lead to (slightly) decreased levels of splicing factors followed by reduced splicing efficiency and splicing defects in specific cell types of patients. Neural crest cells have been postulated to be more sensitive to inefficient splicing than other cell types. Thus, maintenance of splicing factor homeostasis seems to be highly critical for neural crest cellular function and any disturbance in fine-tuning the concentration of specific splicing factors may compromise craniofacial development.[13] The high capability of neural crest cells to autoregulate mutated splicing factor genes is demonstrated by the report of a patient with a heterozygous *SF3B4* nonsense variant who had developmental and growth delay but did not show the typical dysmorphic features of Nager syndrome.[85] Based on these data, we hypothesize that any disturbance in establishing a near homeostatic concentration of specific splicing factors leads to alternative splicing defects in distinct cell types of an individual with a pathogenic variant in a splicing factor gene and a non-predictable spectrum of clinical abnormalities.

In conclusion, we report on 9 subjects with de novo heterozygous *PHF5A* variants and a phenotype consisting of developmental delay, pre- and post-natal growth abnormalities, ear abnormalities, and hypospadias. Data from our functional studies using subject-derived fibroblasts with heterozygous *PHF5A* LOF variants suggest the existence of feedback mechanisms to maintain (almost) normal levels of PHF5A and other SF3B components. These compensatory mechanisms in subject fibroblasts suggest disturbed autoregulation of mutated splicing factor genes in specific cell types, that is, neural crest cells, during embryonic development rather than haploinsufficiency as pathomechanism.

Data Availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplemental Material file. The complete phenotype of included subjects is available as a Phenopacket in the Supplemental Data. Consent restrictions preclude sharing of full data sets and the consents do not cover the deposition of the exome sequencing/genome sequencing data in a public database. The *PHF5A* variants described in the subjects of this manuscript have been submitted to LOVD: #0000305532, #0000902940, #0000902942, #0000902968, #0000902970, #0000902971, #0000903162, #0000908850, and #0000927485.

Acknowledgments

The authors would like to thank all subjects and families for their participation in this study and Jane Rehberg, Henrike Wilshusen, and Dennis Zorndt for their skillful technical assistance.

Funding

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (KU 1240/13-1 to K.K.), the Dutch Organization for Health Research and Development (ZonMw grant 912-12-109 to B.B.A.d.V.), Donders Junior researcher grant 2019 (to B.B.A.d.V.), The Netherlands Organization for Scientific Research (ZonMw Veni, grant 91617021; ZonMw Vidi, grant 09150172110002 to T.S.B.), an Erasmus MC Fellowship 2017 (to T.S.B.), and Erasmus MC Human Disease Model Award 2018 (to T.S.B.). UDP-Vic acknowledges financial support from the Murdoch Children's Research Institute and the Harbig Foundation. The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. Sequencing and analysis were provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung, and Blood Institute grant UM1 HG008900 and in part by National Human Genome Research Institute grant R01 HG009141.

This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome (grant number WT223718/Z/21/Z).

Author Information

Conceptualization: F.L.H., B.B.A.d.V., K.K.; Data curation: F.L.H., A.J.M.D., M.H., R.P., T.B., C.C., C.M., J.-M.F.N., J.F., A.J., C.H., S.M., E.A., A.N., J.v.d.S., R.E., E.v.B., G.M.S.M., M.v.S., T.S.B., E.L.W., A.K., L.D., L.P., S.M.W., B.B.A.d.V.; Formal Analysis: F.L.H., A.J.M.D., B.B.A.d.V., K.K.; Supervision: K.K.; Writing-original draft: F.L.H., A.J.M.D., M.H., R.P., T.B., C.C., C.M., J.-M.F.N., J.F., A.J., C.H., S.M., E.A., A.N., J.v.d.S., R.E., E.v.B., G.M.S.M., M.v.S., T.S.B., E.L.W., A.K., L.D., L.P., S.M.W., B.B.A.d.V., K.K.; Writing-review and editing: F.L.H., B.B.A.d.V., K.K.

Ethics Declaration

Informed consent for genetic analyses was obtained for all affected individuals, and genetic studies were performed as approved by the Ethics Committee of the Hamburg Medical Chamber (PV7038-4438-BO-ff; Hamburg, Germany) and Institutional Review Boards of all other institutions. The parents of the affected individuals provided written informed consent for participation in the study, clinical data and specimen collection, genetic analysis, and publication of relevant findings. Permission to publish photographs was provided for individuals shown in Figure 1B. All clinical data are de-identified.

Conflict of Interest

The authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2023.100927) contains supplemental material, which is available to authorized users.

Affiliations

[1]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [2]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Institute of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany; [4]Bioinformatics Core, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [5]Department of Neurology, Elisabeth-TweeSteden Hospital, Tilburg, The Netherlands; [6]Institute for Clinical Genetics, University Hospital Carl Gustav Carus at the Technische Universität Dresden, Dresden, Germany; [7]Department of Neuropaediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; [8]Division of Medical Genetics, Department of Experimental Medicine, San Camillo-Forlanini Hospital, Sapienza University, Rome, Italy; [9]Laboratory of Medical Genetics, Translational Cytogenomics Research Unit, Bambino Gesù Children Hospital, IRCCS, Rome, Italy; [10]Department of Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [11]Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands; [12]Discovery Unit, Department of Clinical

Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands; [13]North East Thames Regional Genetic Service, Great Ormond Street Hospital for Children, NHS Foundation Trust, London, United Kingdom; [14]All Wales Medical Genomics Service/ Pennaeth Labordy Genomeg Cymru Gyfan, University Hospital of Wales, Heath Park, Cardiff, United Kingdom; [15]Victorian Clinical Genetics Service, Murdoch Children's Research Institute, VIC; [16]Department of Paediatrics, University of Melbourne, Melbourne, Australia; [17]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA

References

1. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell*. 2009;136(4):701-718. http://doi.org/10.1016/j.cell.2009.02.009
2. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol*. 2011;3(7):a003707. http://doi.org/10.1101/cshperspect.a003707
3. Shi Y. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol*. 2017;18(11):655-670. http://doi.org/10.1038/nrm.2017.86
4. Larsen NA. The SF3b complex is an integral component of the spliceosome and targeted by natural product-based inhibitors. *Subcell Biochem*. 2021;96:409-432. http://doi.org/10.1007/978-3-030-58971-4_12
5. Cretu C, Schmitzová J, Ponce-Salvatierra A, et al. Molecular architecture of SF3b and structural consequences of its cancer-related mutations. *Mol Cell*. 2016;64(2):307-319. http://doi.org/10.1016/j.molcel.2016.08.036
6. Golas MM, Sander B, Will CL, Lührmann R, Stark H. Molecular architecture of the multiprotein splicing factor SF3b. *Science*. 2003;300(5621):980-984. http://doi.org/10.1126/science.1084155
7. Will CL, Urlaub H, Achsel T, Gentzel M, Wilm M, Lührmann R. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J*. 2002;21(18):4978-4988. http://doi.org/10.1093/emboj/cdf480
8. Trappe R, Ahmed M, Gläser B, et al. Identification and characterization of a novel murine multigene family containing a PhD-finger-like motif. *Biochem Biophys Res Commun*. 2002;293(2):816-826. http://doi.org/10.1016/S0006-291X(02)00277-2
9. Butt H, Bazin J, Alshareef S, et al. Overlapping roles of spliceosomal components SF3B1 and PHF5A in rice splicing regulation. *Commun Biol*. 2021;4(1):529. http://doi.org/10.1038/s42003-021-02051-y
10. Teng T, Tsai JH, Puyang X, et al. Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A-SF3b complex. *Nat Commun*. 2017;8:15522. http://doi.org/10.1038/ncomms15522
11. Griffin C, Saint-Jeannet JP. Spliceosomopathies: diseases and mechanisms. *Dev Dyn*. 2020;249(9):1038-1046. http://doi.org/10.1002/dvdy.214
12. Jenkins JL, Kielkopf CL. Splicing factor mutations in myelodysplasias: insights from spliceosome structures. *Trends Genet*. 2017;33(5):336-348. http://doi.org/10.1016/j.tig.2017.03.001
13. Beauchamp MC, Alam SS, Kumar S, Jerome-Majewska LA. Spliceosomopathies and neurocristopathies: two sides of the same coin? *Dev Dyn*. 2020;249(8):924-945. http://doi.org/10.1002/dvdy.183
14. Lynch DC, Revil T, Schwartzentruber J, et al. Disrupted autoregulation of the spliceosomal gene SNRPB causes cerebro-costo-mandibular syndrome. *Nat Commun*. 2014;5:4483. http://doi.org/10.1038/ncomms5483
15. Duker A, Velasco D, Robertson N, Jackson A, DeFelice M, Bober M. RNU4atac-opathy. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. *GeneReviews((R))*. Seattle, (WA); 1993. Accessed April 20, 2023. https://www.ncbi.nlm.nih.gov/books/NBK589232/
16. Edery P, Marcaillou C, Sahbatou M, et al. Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science*. 2011;332(6026):240-243. http://doi.org/10.1126/science.1202205
17. Farach LS, Little ME, Duker AL, et al. The expanding phenotype of RNU4ATAC pathogenic variants to Lowry Wood syndrome. *Am J Med Genet A*. 2018;176(2):465-469. http://doi.org/10.1002/ajmg.a.38581
18. He H, Liyanarachchi S, Akagi K, et al. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science*. 2011;332(6026):238-240. http://doi.org/10.1126/science.1200587
19. Merico D, Roifman M, Braunschweig U, et al. Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. *Nat Commun*. 2015;6:8718. http://doi.org/10.1038/ncomms9718
20. Dauber A, Golzio C, Guenot C, et al. SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. *Am J Hum Genet*. 2013;93(5):798-811. http://doi.org/10.1016/j.ajhg.2013.09.010
21. Verheij JB, de Munnik SA, Dijkhuizen T, et al. An 8.35 Mb overlapping interstitial deletion of 8q24 in two patients with coloboma, congenital heart defect, limb abnormalities, psychomotor retardation and convulsions. *Eur J Med Genet*. 2009;52(5):353-357. http://doi.org/10.1016/j.ejmg.2009.05.006
22. Lines M, Hartley T, MacDonald SK, Boycott KM. Mandibulofacial dysostosis with microcephaly. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. *GeneReviews((R))*. Seattle, (WA); 1993. Accessed April 20, 2023. https://www.ncbi.nlm.nih.gov/books/NBK214367/
23. Lines MA, Huang L, Schwartzentruber J, et al. Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes mandibulofacial dysostosis with microcephaly. *Am J Hum Genet*. 2012;90(2):369-377. http://doi.org/10.1016/j.ajhg.2011.12.023
24. Favaro FP, Alvizi L, Zechi-Ceide RM, et al. A noncoding expansion in EIF4A3 causes Richieri-Costa-Pereira syndrome, a craniofacial disorder associated with limb defects. *Am J Hum Genet*. 2014;94(1):120-128. http://doi.org/10.1016/j.ajhg.2013.11.020
25. Ludecke HJ, Wieczorek D. TXNL4A-related craniofacial disorders. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. *GeneReviews((R))*. Seattle, (WA); 1993. Accessed April 20, 2023. https://www.ncbi.nlm.nih.gov/books/NBK373577/
26. Wieczorek D, Newman WG, Wieland T, et al. Compound heterozygosity of low-frequency promoter deletions and rare loss-of-function mutations in TXNL4A causes Burn-McKeown syndrome. *Am J Hum Genet*. 2014;95(6):698-707. http://doi.org/10.1016/j.ajhg.2014.10.014
27. Xu M, Xie YA, Abouzeid H, et al. Mutations in the spliceosome component CWC27 cause retinal degeneration with or without additional developmental anomalies. *Am J Hum Genet*. 2017;100(4):592-604. http://doi.org/10.1016/j.ajhg.2017.02.008
28. El Chehadeh S, Kerstjens-Frederikse WS, Thevenon J, et al. Dominant variants in the splicing factor PUF60 cause a recognizable syndrome with intellectual disability, heart defects and short stature. *Eur J Hum Genet*. 2016;25(1):43-51. http://doi.org/10.1038/ejhg.2016.133
29. Bernier FP, Caluseriu O, Ng S, et al. Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. *Am J Hum Genet*. 2012;90(5):925-933. http://doi.org/10.1016/j.ajhg.2012.04.004
30. McPherson E, Zaleski C, Ye Z, Lin S. Rodriguez syndrome with SF3B4 mutation: a severe form of Nager syndrome? *Am J Med Genet A*. 2014;164A(7):1841-1845. http://doi.org/10.1002/ajmg.a.36555
31. Cassina M, Cerqua C, Rossi S, et al. A synonymous splicing mutation in the SF3B4 gene segregates in a family with highly variable Nager syndrome. *Eur J Hum Genet*. 2017;25(3):371-375. http://doi.org/10.1038/ejhg.2016.176

32. Czeschik JC, Voigt C, Alanay Y, et al. Clinical and mutation data in 12 patients with the clinical diagnosis of Nager syndrome. *Hum Genet*. 2013;132(8):885-898. http://doi.org/10.1007/s00439-013-1295-2
33. Drivas TG, Taylor JA, Zackai EH. The final demise of Rodriguez lethal acrofacial dysostosis: a case report and review of the literature. *Am J Med Genet A*. 2019;179(6):1063-1068. http://doi.org/10.1002/ajmg.a.61121
34. Irving MD, Dimitrov BI, Wessels M, Holder-Espinasse M, Chitayat D, Simpson MA. Rodriguez acrofacial dysostosis is caused by apparently de novo heterozygous mutations in the SF3B4 gene. *Am J Med Genet A*. 2016;170(12):3133-3137. http://doi.org/10.1002/ajmg.a.37946
35. Lund IC, Vestergaard EM, Christensen R, Uldbjerg N, Becher N. Prenatal diagnosis of Nager syndrome in a 12-week-old fetus with a whole gene deletion of SF3B4 by chromosomal microarray. *Eur J Med Genet*. 2016;59(1):48-51. http://doi.org/10.1016/j.ejmg.2015.12.001
36. Petit F, Escande F, Jourdain AS, et al. Nager syndrome: confirmation of SF3B4 haploinsufficiency as the major cause. *Clin Genet*. 2014;86(3):246-251. http://doi.org/10.1111/cge.12259
37. Timberlake AT, Griffin C, Heike CL, et al. Haploinsufficiency of SF3B2 causes craniofacial microsomia. *Nat Commun*. 2021;12(1):4680. http://doi.org/10.1038/s41467-021-24852-9
38. Yang M, Lu X, Zhang Y, et al. Whole-exome sequencing analysis in 10 families of sporadic microtia with thoracic deformities. *Mol Genet Genomic Med*. 2021;9(5):e1657. http://doi.org/10.1002/mgg3.1657
39. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet*. 2008;83(5):610-615. http://doi.org/10.1016/j.ajhg.2008.09.017
40. Jacobsen JOB, Baudis M, Baynam GS, et al. The GA4GH Phenopacket schema defines a computable representation of clinical data. *Nat Biotechnol*. 2022;40(6):817-820. http://doi.org/10.1038/s41587-022-01357-4
41. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890. http://doi.org/10.1093/bioinformatics/bty560
42. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. http://doi.org/10.1093/bioinformatics/bts635
43. von Elsner L, Chai G, Schneeberger PE, et al. Biallelic FRA10AC1 variants cause a neurodevelopmental disorder with growth retardation. *Brain*. 2022;145(4):1551-1563. http://doi.org/10.1093/brain/awab403
44. Trincado JL, Entizne JC, Hysenaj G, et al. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol*. 2018;19(1):40. http://doi.org/10.1186/s13059-018-1417-1
45. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14(4):417-419. http://doi.org/10.1038/nmeth.4197
46. Li YI, Knowles DA, Humphrey J, et al. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet*. 2018;50(1):151-158. http://doi.org/10.1038/s41588-017-0004-9
47. Brechtmann F, Mertes C, Matusevičiūtė A, et al. OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am J Hum Genet*. 2018;103(6):907-917. http://doi.org/10.1016/j.ajhg.2018.10.025
48. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. http://doi.org/10.1186/s13059-014-0550-8
49. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. 2015;519(7542):223-228. http://doi.org/10.1038/nature14135
50. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433-438. http://doi.org/10.1038/nature21062
51. Sobreira N, Schiettecatte F, Boehm C, Valle D, Hamosh A. New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. *Hum Mutat*. 2015;36(4):425-431. http://doi.org/10.1002/humu.22769
52. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612. http://doi.org/10.1002/jcc.20084
53. Dunbrack RL Jr. Rotamer libraries in the 21st century. *Curr Opin Struct Biol*. 2002;12(4):431-440. http://doi.org/10.1016/s0959-440x(02)00344-5
54. Kiniry SJ, Judge CE, Michel AM, Baranov PV. Trips-viz: an environment for the analysis of public and user-generated ribosome profiling data. *Nucleic Acids Res*. 2021;49(W1):W662-W670. http://doi.org/10.1093/nar/gkab323
55. Tirosh O, Cohen Y, Shitrit A, et al. The transcription and translation landscapes during human Cytomegalovirus infection reveal novel Host-Pathogen Interactions. *PLoS Pathog*. 2015;11(11):e1005288. http://doi.org/10.1371/journal.ppat.1005288
56. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res*. 2001;29(5):1185-1190. http://doi.org/10.1093/nar/29.5.1185
57. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol*. 1997;4(3):311-323. http://doi.org/10.1089/cmb.1997.4.311
58. Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res*. 1987;15(17):7155-7174. http://doi.org/10.1093/nar/15.17.7155
59. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol*. 2004;11(2-3):377-394. http://doi.org/10.1089/1066527041410418
60. Zhang MQ. Statistical features of human exons and their flanking regions. *Hum Mol Genet*. 1998;7(5):919-932. http://doi.org/10.1093/hmg/7.5.919
61. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat*. 2019;40(8):1030-1038. http://doi.org/10.1002/humu.23798
62. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. http://doi.org/10.1038/s41586-020-2308-7
63. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291. http://doi.org/10.1038/nature19057
64. Coban-Akdemir Z, White JJ, Song X, et al. Identifying genes whose mutant transcripts cause dominant disease traits by potential gain-of-function alleles. *Am J Hum Genet*. 2018;103(2):171-187. http://doi.org/10.1016/j.ajhg.2018.06.009
65. Xie Z, Bailey A, Kuleshov MV, et al. Gene set knowledge discovery with Enrichr. *Curr Protoc*. 2021;1(3):e90. http://doi.org/10.1002/cpz1.90
66. Gao J, Yang D, Cao R, et al. The role of Fbxo5 in the development of human malignant tumors. *Am J Cancer Res*. 2022;12(4):1456-1464. https://pubmed.ncbi.nlm.nih.gov/35530293/
67. Hao Z, Zhang H, Cowell J. Ubiquitin-conjugating enzyme UBE2C: molecular biology, role in tumorigenesis, and potential as a biomarker. *Tumour Biol*. 2012;33(3):723-730. http://doi.org/10.1007/s13277-011-0291-1
68. Ishimi Y. Regulation of MCM2-7 function. *Genes Genet Syst*. 2018;93(4):125-133. http://doi.org/10.1266/ggs.18-00026
69. Loukil A, Cheung CT, Bendris N, Lemmers B, Peter M, Blanchard JM. Cyclin A2: at the crossroads of cell cycle and cell invasion. *World J Biol Chem*. 2015;6(4):346-350. http://doi.org/10.4331/wjbc.v6.i4.346
70. Vagnarelli P. Repo-man at the intersection of chromatin remodelling, DNA repair, nuclear envelope organization, and cancer progression. *Adv Exp Med Biol*. 2014;773:401-414. http://doi.org/10.1007/978-1-4899-8032-8_18
71. Marques F, Tenney J, Duran I, et al. Altered mRNA splicing, chondrocyte gene expression and abnormal skeletal development due to SF3B4 mutations in Rodriguez acrofacial dysostosis. *PLoS Genet*. 2016;12(9):e1006307. http://doi.org/10.1371/journal.pgen.1006307

72. Jumaa H, Nielsen PJ. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J*. 1997;16(16):5077-5085. http://doi.org/10.1093/emboj/16.16.5077
73. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature*. 2007;446(7138):926-929. http://doi.org/10.1038/nature05676
74. Ni JZ, Grate L, Donohue JP, et al. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev*. 2007;21(6):708-718. http://doi.org/10.1101/gad.1525507
75. Saltzman AL, Kim YK, Pan Q, Fagnani MM, Maquat LE, Blencowe BJ. Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol Cell Biol*. 2008;28(13):4320-4330. http://doi.org/10.1128/MCB.00361-08
76. Saltzman AL, Pan Q, Blencowe BJ. Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev*. 2011;25(4):373-384. http://doi.org/10.1101/gad.2004811
77. Sureau A, Gattoni R, Dooghe Y, Stévenin J, Soret JSC. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J*. 2001;20(7):1785-1796. http://doi.org/10.1093/emboj/20.7.1785
78. Ding F, Su CJ, Edmonds KK, Liang G, Elowitz MB. Dynamics and functional roles of splicing factor autoregulation. *Cell Rep*. 2022;39(12):110985. http://doi.org/10.1016/j.celrep.2022.110985
79. Sun S, Zhang Z, Sinha R, Karni R, Krainer AR. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nat Struct Mol Biol*. 2010;17(3):306-312. http://doi.org/10.1038/nsmb.1750
80. Wu H, Sun S, Tu K, et al. A splicing-independent function of SF2/ASF in microRNA processing. *Mol Cell*. 2010;38(1):67-77. http://doi.org/10.1016/j.molcel.2010.02.021
81. Beauchamp MC, Djedid A, Bareke E, et al. Mutation in Eftud2 causes craniofacial defects in mice via mis-splicing of Mdm2 and increased P53. *Hum Mol Genet*. 2021;30(9):739-757. http://doi.org/10.1093/hmg/ddab051
82. Devotta A, Juraver-Geslin H, Gonzalez JA, Hong CS, Saint-Jeannet JP. Sf3b4-depleted Xenopus embryos: a model to study the pathogenesis of craniofacial defects in Nager syndrome. *Dev Biol*. 2016;415(2):371-382. http://doi.org/10.1016/j.ydbio.2016.02.010
83. Lei L, Yan SY, Yang R, et al. Spliceosomal protein eftud2 mutation leads to p53-dependent apoptosis in zebrafish neural progenitors. *Nucleic Acids Res*. 2017;45(6):3422-3436. http://doi.org/10.1093/nar/gkw1043
84. Park BY, Tachi-Duprat M, Ihewulezi C, Devotta A, Saint-Jeannet JP. The core splicing factors EFTUD2, SNRPB and TXNL4A are essential for neural crest and craniofacial development. *J Dev Biol*. 2022;10(3):29. http://doi.org/10.3390/jdb10030029
85. Cadieux-Dion M, Hughes S, Engleman K, Rush ET, Saunders C. Nager syndrome in patient lacking acrofacial dysostosis: expanding the phenotypic spectrum of SF3B4-related disease. *Am J Med Genet A*. 2021;185(5):1515-1518. http://doi.org/10.1002/ajmg.a.62113

Update

Genetics in Medicine
Volume 25, Issue 11, November 2023, Page

DOI: https://doi.org/10.1016/j.gim.2023.100964

Genetics in Medicine (2023) 25, 100964

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

CORRECTION

De novo *PHF5A* variants are associated with craniofacial abnormalities, developmental delay, and hypospadias

Frederike L. Harms[1], Alexander J.M. Dingemans[2], Maja Hempel[1,3], Rolph Pfundt[2], Tatjana Bierhals[1], Christian Casar[4], Christian Müller[4], Jikke-Mien F. Niermeijer[5], Jan Fischer[6], Arne Jahn[6], Christoph Hübner[7], Silvia Majore[8], Emanuele Agolini[9], Antonio Novelli[9], Jasper van der Smagt[10], Robert Ernst[10], Ellen van Binsbergen[10], Grazia M.S. Mancini[11], Marjon van Slegtenhorst[11], Tahsin Stefan Barakat[11,12], Emma L. Wakeling[13], Arveen Kamath[14], Lilian Downie[15,16], Lynn Pais[17], Susan M. White[15,16], Bert B.A. de Vries[2], Kerstin Kutsche[1]

[1]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [2]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Institute of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany; [4]Bioinformatics Core, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [5]Department of Neurology, Elisabeth-TweeSteden Hospital, Tilburg, The Netherlands; [6]Institute for Clinical Genetics, University Hospital Carl Gustav Carus at the Technische Universität Dresden, Dresden, Germany; [7]Department of Neuropaediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; [8]Division of Medical Genetics, Department of Experimental Medicine, San Camillo-Forlanini Hospital, Sapienza University, Rome, Italy; [9]Laboratory of Medical Genetics, Translational Cytogenomics Research Unit, Bambino Gesù Children Hospital, IRCCS, Rome, Italy; [10]Department of Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [11]Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands; [12]Discovery Unit, Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands; [13]North East Thames Regional Genetic Service, Great Ormond Street Hospital for Children, NHS Foundation Trust, London, United Kingdom; [14]All Wales Medical Genomics Service/ Pennaeth Labordy Genomeg Cymru Gyfan, University Hospital of Wales, Heath Park, Cardiff, United Kingdom; [15]Victorian Clinical Genetics Service, Murdoch Children's Research Institute, Victoria, Australia; [16]Department of Paediatrics, University of Melbourne, Melbourne, Australia; [17]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA

Correction to: *Genetics in Medicine* 2023; https://doi.org/10.1016/j.gim.2023.100927, published online 6 July 2023.

In the article "De novo *PHF5A* variants are associated with craniofacial abnormalities, developmental delay, and hypospadias" (*Genet Med* 2023;25:100927), the following update was made. In the author listing, the author's name "Tahshin S. Barakat" has been updated to "Tahsin Stefan Barakat." Please see revised author listing shown above. The article has been corrected online can be accessed at https://doi.org/10.1016/j.gim.2023.100927.

doi: https://doi.org/10.1016/j.gim.2023.100964