Mutations in *TAF8* cause a neurodegenerative disorder

Keit Men Wong[1,2], Wayne M. Jepsen[3], Stephanie Efthymiou[4], Vincenzo Salpietro[4], Meredith Sanchez-Castillo[3], Janice Yip[4], Yamna Kriouile[5], Susann Diegmann[6], Steffi Dreha-Kulaczewski[6], Janine Altmüller[7], Holger Thiele[8], Peter Nürnberg[8], Mehran Beiraghi Toosi[9], Javad Akhondian[9], Ehsan Ghayoor Karimiani[10], Hannah Hummel-Abmeier[6], Brenda Huppke[1,2], Henry Houlden[4], Jutta Gärtner[6], Reza Maroofian[4], Peter Huppke[1,2]

Abstract

TAF8 is part of the transcription factor II D (TFIID) complex, composed of the TATA-binding protein (TBP) and 13 TBP-associated factors (TAFs). TFIID is the first general transcription factor recruited at promoters to assemble the RNA polymerase II preinitiation complex. So far disorders related to variants in five of the 13 subunits of human TFIID have been described. Recently, a child with a homozygous c.781–1G>A mutation in *TAF8* has been reported. Here we describe seven further patients with mutations in *TAF8* and thereby confirm the *TAF8* related disorder. In two sibling patients, we identified two novel compound heterozygous *TAF8* splice site mutations, c.45+4A>G and c.489G>A, which cause aberrant splicing as well as reduced expression and mislocalization of TAF8. In five further patients, the previously described c.781–1G>A mutation was present on both alleles. The clinical phenotype associated with the different *TAF8* mutations is characterized by severe psychomotor retardation with almost absent development, feeding problems, microcephaly, growth retardation, spasticity and epilepsy. Cerebral imaging showed hypomyelination, a thin corpus callosum and brain atrophy. Moreover, repeated imaging in the sibling pair demonstrated progressive cerebral and cerebellar atrophy. Consistently, reduced N-acetylaspartate, a marker of neuronal viability, was observed on magnetic resonance spectroscopy. Further review of the literature shows that mutations causing a reduced expression of TFIID subunits have an overlapping phenotype of microcephaly, developmental delay and intellectual disability. Although TFIID plays an important role in RNA polymerase II transcription in all cells and tissues, the symptoms associated with such defects are almost exclusively neurological. This might indicate a specific vulnerability of neuronal tissue to widespread deregulation of gene expression as also seen in Rett syndrome or Cornelia de Lange syndrome.

Author affiliations:

1 Department of Neuropediatrics, Jena University Hospital, 07747 Jena, Germany

2 Center for Rare Diseases, Jena University Hospital, 07747 Jena, Germany

3 Translational Genomics Research Institute, Neurogenomics Division - Huentelman Lab,. Phoenix, AZ 85004, US

4 Department of Neuromuscular Disorders, UCL Institute of Neurology, London WC1N 3BG, UK

5 Unit of Neuropediatrics and Neurometabolism, Pediatric Department 2, Rabat Children's Hospital, BP 6527 Rabat, Morocco.

6 Department of Pediatrics and Pediatric Neurology, University Medical Center Göttingen, 37075 Goettingen, Germany.

7 Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Core Facility Genomics, Charitéplatz 1, 10117 Berlin, Germany. Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 10117 Berlin, Germany

8 Cologne Center for Genomics (CCG) and Center for Molecular Medicine Cologne (CMMC), University of Cologne, Faculty of Medicine and University Hospital Cologne, 50931 Cologne, Germany.

9 Department of Pediatric Neurology, Ghaem Hospital, Mashhad University of Medical Sciences, Mashhad 45785, Iran

10 Molecular and Clinical Sciences Institute, St. George's, University of London, Cranmer Terrace, London SW17 0RE, UK, Innovative medical research center, Mashhad branch, Islamic Azad University, Mashhad 45785, Iran

Correspondence to: Peter Huppke

Department of Neuropediatrics, Jena University Hospital, Jena, Germany

Am Klinikum 1, 07747 Jena

Tel. +49 3641/9 329-651 Fax 49 3641/9 329-652

Email: phuppke@med.uni-goettingen.de

Running title: Mutations in TAF8 cause a neurodegenerative disorder

Keywords: TAF8; TFIID; neurodegenerative; psychomotor retardation

Abbreviations: CE=cytoplasmic extract; Cho=choline-containing compounds; FBS=fetal bovine serum; GM=gray matter; HCF=Head circumference; Ins=myo-inositol; mESCs=mouse embryonic stem cells; MRS= magnetic resonance spectroscopy; NE=nuclear extraction; Pol II = polymerase II; SCA17=Spinocerebellar Ataxia 17; TAFs =TBP-associated factors; TBP= TATA-binding protein; tCr=creatine and phosphocreatine; TFIID= transcription factor II D; tNAA=N-acetylaspartate and N-acetylaspartylglutamate; XDP=X-Linked Dystonia-Parkinsonism

Introduction

RNA polymerase II (Pol II) transcribes all mRNA genes in eukaryotes. To initiate transcription by Pol II, a transcriptional preinitiation complex must assemble at the promotor. The first step is the binding of the general transcription factor TFIID, followed by other general transcription factors and Pol II. Human TFIID consists of the TATA-binding protein (TBP) and 13 TBP-associated factors (TAFs)[1]. The number of genetic disorders associated with variants in the genes coding for subunits of TFIID is growing. A polyQ expansion encoded by a tandem repeat of CAG trinucleotides in the *TBP* gene has been found to cause Spinocerebellar Ataxia 17 (SCA17), a dominant neurodegenerative disorder characterized by ataxia, dystonia, parkinsonism and dementia usually starting at middle age[2]. As for other polyQ diseases, the mutation primarily causes a gain of function, related to the propensity of the mutated protein to form aggregates[3]. Subtelomeric deletions of the long arm of chromosome 6 that include the *TBP* gene, however, are associated with intellectual disability, developmental delay, dysmorphic features, seizures, hypotonia, microcephaly and hypoplasia of the corpus callosum[4]. The first human disorder associated with variants in a gene encoding one of the TAF proteins was X-Linked Dystonia-Parkinsonism (XDP)[5]. Here, an intronic SINE-VNTR-Alu (SVA)-type retrotransposon results in aberrant truncated transcripts of *TAF1*[6]. The disorder predominantly affects Filipino men in their fifth decade of life and is characterized by progressive dystonia, often accompanied by parkinsonism. Survival in most patients is only 10 -12 years[7]. Intriguingly, missense mutations in the same gene have been described in patients with hypotonia, facial dysmorphism, and developmental delay[8]. More recently, neurological disorders related to variants in *TAF2, TAF6* and *TAF13* have also been described. Autosomal

recessive mutations in *TAF2* in 3 families have been associated with postnatal microcephaly, pyramidal signs, thin corpus callosum, and intellectual disability[9, 10]. In addition, six patients from 3 families with homozygous missense mutations in *TAF6* presented with moderate intellectual disability, nystagmus, mild dysmorphic features and stereotypic behavior[11-13]. Two of the identified mutations reportedly disrupt the interactions between TAF6 and other TFIID components[13].Also reported are four patients from two families with homozygous missense mutations in *TAF13* and a phenotype of mild intellectual disability, microcephaly and severe growth retardation[14]. Both detected missense mutations impair the formation of a histone-like heterodimer of TAF13 -TAF11 leading to deregulation of gene expression. Finally, in 2018, a single patient with a homozygous splice site mutation in *TAF8* was described[15]. In this article, we describe the molecular and clinical phenotype of 7 further patients with mutations in *TAF8*.

Materials and methods

MR spectroscopy

Localized proton MR spectrum (64 accumulations) was acquired at a 3 T clinical scanner (Magnetom TIM Trio, Siemens Healthcare, Erlangen, Germany) using a STEAM localization sequence with repetition time (TR)/echo time (TE)/mixing time (TM)= 6000/20/10 ms as described[16, 17]. Volume-of-interest (12,5 ml) was placed within posterior cortical gray matter (GM). The age matched healthy control spectrum was selected from our local database and displayed in Fig. 2N. The important metabolites for this study comprise the neuroaxonal markers N-acetylaspartate and N-acetylaspartylglutamate (tNAA) which resonate in one prominent peak at 2.01 ppm (Fig. 2M, N). Furthermore, creatine and phosphocreatine (tCr) as ubiquitous compounds linked to energy metabolism, choline-containing compounds (Cho) involved in membrane turnover, and the glial marker myo-inositol (Ins) are detectable and qualitatively compared to the control.

Whole exome sequencing

Patients 1 and 2: DNA samples were obtained and extracted from EDTA blood after informed consent and approval by the ethic commission from the University of Medical Center Göttingen, Göttingen (approval number 2516). Barcoded library enrichment for multiplex whole exome sequencing was performed using the Agilent SureSelect Human All Exon V6 kit (60 Mb) (Agilent Technologies, Santa Clara, USA). Resulting purified and quantified library

pools were sequenced with an allocation of 6 exomes per lane on an Illumina Hiseq 4000 system (Illumina, Inc., San Diego, USA) using 2x76 bp paired-end protocol. Following data processing, variant calling, annotation and filtering were performed using the ‚Varbank 2' GUI and pipeline version 3.3 (CCG, University of Cologne, Germany). Read alignment to the human genome reference build GRCh38 was performed using the BWA-MEM alignment algorithm. GATK HaplotypeCaller, Samtools mpileup and Platypus variation calls were filtered for high-quality (SNV: QD≥5; ARF≥0.25; MQ≥50; FS≤40; MQRankSum≥-5; ReadPosRankSum≥-5; InDels: QD≥4; FS≤100; ReadPosRankSum≥-5; passed VQSR filter) rare (MAF ≤ 0.01 based on the maximum observed population allele frequency in gnomAD v2) variants, predicted to modify a protein sequence or to impair splicing, implicated by reduced maximum entropy scores (MaxEntScan)[18-20]. Pipeline related false positive variants were further reduced by taking advantage of our InHouseDB. Dataset of the two siblings considering the healthy parents was filtered for rare recessive genetic variants common in both patients to detect their common genetic cause. Filtering was focused on affected genes by two or more heterozygous variants or homozygous variants. In addition, the dataset was checked for variants in genes known to be regulated by genomic imprinting based on the Catalogue of Imprinted Genes www.otago.ac.nz/IGC. The pathogenicity of all received variants was evaluated using several in silico analysis tools by the dbNSFP/dbscSNV version 3.4 database. Validation of all variants of interest and the segregation analysis within the family was performed using Sanger Sequencing.

Patient 3: DNA samples from the proband and both biological parents were extracted from whole blood after informed consent and approval by the Western Institutional Review Board (Protocol #20120789). PCR free whole genome sequencing was performed at NovoGene (Beijing, China) using the NEBNext Ultra II DNA Kit (New England Biolabs, Ipswich, MA, USA) and the NovaSeq 6000 platform (Illumina, San Diego, CA, USA). Reads were aligned to the human genome (Hg19/GRC37) using the Burrows-Wheeler Aligner (BWA-MEM v.0.7.8)(arXiv:1303.3997v2 [q-bio.GN]). PCR duplicates were identified using Picard MarkDuplicates v1.79, and base quality recalibration and indel realignment were performed using the Genome Analysis Toolkit (GATK v3.5-1)[21]. Variants were joint-called with HaplotypeCaller and then recalibrated with GATK (bioRxiv, 201178. DOI: 10.1101/201178). Quality assessments were conducted using FASTQC v0.11.5. Called variants were annotated with SnpEff v3.0a[22] against Ensembl GRCh37.66 and filtered against dbSNP137 1000 Genomes Project (minor allele frequency < 0.05), SnpEff Impact: High + Moderate, GATK

quality score > 300, and an in-house database of highly-variable genes. Both parents were found to be carriers, while the proband was found to be homozygous, for a ClinVar pathogenic *TAF8* splice-site variant (NM_138572.2, c.781-1G>A) that phenotypically matches the autosomal recessive disorder previously described[15]. Sanger sequencing was performed at GeneDx (Gaithersburg, MD, USA) for validation of the variant and zygosity.

Patients 4-7: This study was approved by local institutional IRB/ethical review boards, and written informed consent was obtained prior to genetic testing from the families involved. Genomic DNA was extracted from peripheral blood samples according to standard procedures of phenol chloroform extraction. WES on each proband was performed as described elsewhere [23]. Briefly, target enrichment was performed with 2 μg genomic DNA using the SureSelectXT Human All Exon Kit version 6 (Agilent Technologies, Santa Clara, CA, USA) to generate barcoded whole-exome sequencing libraries. Libraries were sequenced on the HiSeqX platform (Illumina, San Diego, CA, USA) with 50x coverage. Quality assessment of the sequence reads was performed by generating QC statistics with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc).
Our bioinformatics filtering strategy included screening for only exonic and donor/acceptor splicing variants. In accordance with the pedigree and phenotype, priority was given to rare variants (<0.01% in public databases, including 1,000 Genomes project, NHLBI Exome Variant Server, Complete Genomics 69, and Exome Aggregation Consortium [ExAC v0.2]) that were fitting a recessive (homozygous or compound heterozygous) or a de novo model and/or variants in genes previously linked to developmental delay, intellectual disability and other neurological disorders.

Cell culture

Fibroblasts were maintained as monolayer cultures in Dulbecco's modified Eagle's medium (DMEM/low glucose) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine and 100 U/ml penicillin, 100 μg/ml streptomycin. Cells were incubated at 37 °C in an atmosphere of 5% CO2.

Polymerase chain reaction and Sanger sequencing

Polymerase chain reaction (PCR) of gDNA and cDNA sequences flanking splice site c.45+4A>G variant and splice site c.489G>A variant of *TAF8* were performed using specific

primer sets. Detailed primer information is available in Supplementary Table 4. PCR or RT-PCR products, where the specific bands were eluted from the gel using NucleoSpin® Gel and PCR Clean-up kit (Cat# 740609-250, Machery-Nagel, Germany), were subsequently processed for direct dye terminator sequencing with BigDye Terminator Ready Reaction chemistry 3.1 on the Applied Biosystem 3500 genetic analyzer (Thermo Fisher, Germany). All reactions were performed according to provided manufacturer's instructions. The sequences were aligned with the reference sequence of *TAF8* (NCBI: NM_138572.2) gene.

Real time quantitative PCR (RT-qPCR)

Total RNA was extracted from patient and wild-type fibroblasts using NucleoSpin RNA Kit (Macherey-Nagel, Düren, Germany) as described by the manufacturer. RNA quality was verified by gel electrophoresis and OD measurements. cDNA was synthesized using the SuperScript III Reverse Transcriptase with Oligo d (T) primer (Cat#18080-051, Life Technologies) according to manufacturer's instructions. 2µg of total RNA in a final reaction volume of 20µl reaction was converted to cDNA for each condition. The RT–qPCR reactions were performed using QuantStudio3 Real time PCR system (Thermo Fisher Scientific) at annealing temperatures of 60 °C and specificity was controlled by post-amplification melting curve analysis. No template control (NTC) was included in every assay. RT–qPCR quantification was calculated according to the ΔΔCT method[24]. Data were normalized to *ACTB* as reference gene and the mRNA expression levels are expressed relative to the control fibroblasts. Detailed primer information is available in supplementary table 4. All RT-qPCR experiments were carried out on at least three independent replicates.

Western blotting

Cultured fibroblasts were washed with ice cold PBS and lysed in RIPA buffer supplemented with protease inhibitor (Roche) by pipetting. Homogenates were left for 30min on ice, centrifuged at 14,000rpm for 15min at 4°C and the supernatant was prepared for western blotting using standard procedures. Total protein concentration was carried out using BCA protein assay (Interchim, France). Western blotting of cell protein lysates was performed using Bolt Mini gels and Mini Bolt wet transfer modules (Invitrogen) according to the manufacturer's instructions. Blot documentation was performed with a CCD digital imaging system (LAS-3000 Mini, GE Healthcare). Integrated band intensities of protein bands were measured using Image J software and normalized by comparison with the loading control. Primary antibodies

used were: TAF8 (pAb 3478, which generated against the entire TAF8 protein, 1:1000 dilution), TAF8 (Abcam 204894, which detected peptide sequence 85-154, 1:800 dilution) and TAF10 (MABE1079, 1:1000 dilution). Horseradish peroxidase-coupled secondary antibodies from Jackson ImmunoResearch and Lumi-Light and Lumi-Light Plus blotting substrate (Roche, Mannheim, Germany) were used for signal detection.

Immunofluorescence

Fibroblasts were seeded and grown on coverslips. Next day, cells on coverslips were fixed in cold 4% paraformaldehyde for 20 min at RT. Immunofluorescence on fibroblasts was performed using standard procedures. Briefly, after post-fixation, cells were washed three times with cold PBS and permeabilized with 0.05M TBS solution containing 0.5M ammonium chloride and 0.25% Triton X-100 for 10 min. Cells were then washed with 0.05M TBS solution, blocked with 5% normal horse serum in 0.05M TBS for 1hr at RT and incubated with primary antibody (TAF8, Abcam 204894 at 1:200 dilution) overnight at 4°C. After washing, secondary antibody (Jackson Immuno Research Cat# 111-165-144, 1:500 dilution) was applied for 1 h and the coverslips with cells were mounted onto the SuperFrost plus slide (Thermo Fisher Scientific, Germany) with proLong Gold antifade reagent with DAPI (Invitrogen Ref# P36935). The stained cells were analyzed using fluorescent microscopy with Apotome.2 (Zeiss, Germany).

Subcellular fractionation and whole cell extraction

The subcellular fractionation protocol used was adapted from Beg et al. (1993) with modifications[25]. Briefly, fibroblasts were trypsinized, washed with ice-cold PBS and pelleted by centrifugation at 1000 rpm for 5min. Cells were resuspended in cytoplasmic extract (CE) buffer (10mM HEPES, 60mM KCl, 1mM EDTA, 1mM DTT, 1mM PMSF and 0.075% NP40 and supplemented with protease inhibitor; solution is adjusted to pH7.6), vortexed for 10sec and incubated on ice for 10min. Lysate was then spun at 1500 rpm for 4min. The supernatant was carefully removed presenting the cytoplasmic fraction. The remaining nuclear pellet was resuspended in nuclear extraction (NE) buffer (20mM Tris HCl, 420mM NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 1mM PMSF, 25% of glycerol, and supplemented with protease inhibitor; solution is adjusted to pH8) (proportional to 1/2 amount of used CE-buffer). The salt concentration was adjusted by 5M NaCl (proportional to 1/3 amount of used CE-buffer). Then, lysates were sonicated and incubated on ice for 30min. Lastly, the cytoplasmic fraction and

nuclear fraction were spun at maximum speed for 10 min at 4°C to pellet any debris. The resultant supernatants were prepared for western blotting. The whole cell extract was prepared separately, lysed and sonicated in RIPA buffer (400mM NaCl, 50mM Tris-HC1 (pH7.5), 1% NP-40, 0.8% sodium deoxycholate, 0.1% SDS and 5mM EDTA, and supplemented with protease inhibitor). Lysates were left for 30min on ice, then centrifuged at 14,000 rpm for 15min at 4°C, and the supernatant was prepared for western blotting using standard procedures. Total protein concentration was carried out using BCA protein assay (Interchim, France). The nuclear fraction from the control and patient lanes were first normalized to their internal loading control Histone H3. The relative abundance ratio was calculated by dividing the normalized value of the patient by the normalized value of the control. The same normalization was performed for the cytoplasmic fraction, except we used GAPDH as the internal loading control.

Statistical analysis

Statistical analysis of the mRNA expression was performed using the Prism 8 (GraphPad software). Statistical differences were considered to be significant when $p < 0.05$. All quantitative data were presented as mean (± SD.).

Data availability

The derived data that support the findings of this study are available from the corresponding author, on reasonable request.

Results

Clinical characteristics of the affected individuals

The clinical phenotype of the patients with variants in *TAF8,* including the previously published case, is summarized in table 1, the pedigrees in Fig.1.

Patient 1, a Caucasian female, was born at 38 weeks after an uneventful pregnancy to non consanguineous parents. She showed no dysmorphic features at birth. In the first weeks of life poor sucking and swallowing resulted in a 20% loss of birth weight; feeding difficulties have remained an ongoing problem since. When first seen in Göttingen at 16 months of age she showed severe psychomotor retardation; she was able to roll over but was unable to sit or stand,

she did not speak or react to language and did not fixate objects or faces. According to the parents, however, she had followed faces in the first 12 months of life. Head circumference (HCF) was normal at birth (34.5cm) but by age 16 months was below the 3rd centile (43cm) as was height (72cm) and weight (7.2kg). At age 5 years she developed epilepsy with short generalized seizures. Multifocal epileptic discharges were visible on EEG and treatment with levetiracetam resulted in initial control of seizures. When last seen at age 10 years she had progressive microcephaly (HCF 46cm), growth retardation (height 117.5cm) and was underweight (weight 16.3kg). She reacted to noises but did not speak or follow instructions. She showed severe spasticity with contractures and suffered frequent seizures despite anticonvulsive therapy.

Patient 2, the younger sister of patient 1, was also born after a normal pregnancy at 38 weeks without dysmorphic features. Like her sister, she had trouble sucking and swallowing and showed only minimal signs of psychomotor development. At first presentation at age 21 months she did not fixate or follow faces or objects, had no head control and was unable to roll over, sit or stand. HCF, weight and height were below the 3rd centile. At age 5 years, when last seen, her developmental status was unchanged. Generalized spasticity was present but no contractures. She showed poor visual fixation and was unable to grasp or hold toys. She had also developed epilepsy suffering short generalized seizures.

Patient 3, male, was born via C-section due to preeclampsia at 36-weeks to distantly related (eighth cousins) parents of Mexican ancestry. After birth he presented with feeding difficulties, jaundice, hypoxia, and hypoglycemia. Due to ongoing severe dysphagia and gastroesophageal reflux disease, placement of a feeding tube was required. His development was globally delayed with almost no motor development, no initial head control and pronounced axial hypotonia. Although electroretinography and visual evoked potential test results were normal he demonstrated signs of visual impairment with poor visual fixation and focus. At age 18 months he began to roll over and at 2 years and 4 months of age he could lift his head in the prone position and bear weight on his legs but was unable to sit without assistance. He did not grasp or hold toys but could move and smack his hands and kick his legs.

Patient 4, the son of a healthy consanguineous couple (first cousins) of Moroccan origin was born at term after an uneventful pregnancy. Psychomotor retardation was already noted at age 4 months when the patient presented with a swallowing disorder and reduced spontaneous movements. At this time, first epileptic seizures were observed, initially with attacks of cyanosis and later focal clonic seizures of the left lower limb followed by generalized tonic and tonic-clonic seizures unresponsive to sodium valproate. By age 6 years epileptic seizures resolved.

When last seen at 9 years of age, he presented with microcephaly, poor growth and short stature. Neuromotor development was severely delayed; he had not achieved independent sitting (only with support) or walking, was able to move his hands and feet, but unable to grasp objects. Clinical examination showed axial hypotonia and poor neck and truncal control. He had not developed any speech but was able to follow with his gaze, turn his head to noises, smile and recognize family members Serious feeding difficulties with gastroesophageal reflux disease continued to be a major problem.

Patient 5, the older brother of patient 4, was also born at term after an uneventful pregnancy. At birth he presented with a delayed cry, cyanosis and swallowing problems. Like his brother, he showed severe psychomotor retardation with generalized hypotonia, never learned to sit unsupported or walk. He acquired no language but recognized his family. At 2 years of age, he developed epilepsy with generalized tonic-clonic seizures unresponsive to sodium valproate. The seizures resolved at the age of 6 years. Around the age of 10, the patient began to exhibit psychomotor restlessness and self-harm (biting his fingers) and, for a period of 3 years, the patient suffered bouts of crying with intermittent insomnia-like sleep disorders for several days. When last seen at 19 years of age, the patient had microcephaly, poor growth and short stature. He was unable to sit or walk, could not grasp objects and showed axial hypotonia and poor neck and truncal control on clinical examination. The parents reported signs of regression with loss of movement and progressive spasticity of the lower limbs. He showed no speech development but could smile and recognize family members.

Patient 6 and 7 are female siblings. Both were born at term to consanguineous Iranian parents (first cousins). Feeding difficulties and poor growth have been apparent since birth in both patients resulting in short stature and microcephaly. At age 40 and 50 days, respectively, both developed tonic seizures. Both children have shown almost no development and now, aged 5 and 15 years, cannot roll-over, sit walk or talk. Neurological findings since the first year of life include spasticity in the upper and lower limbs, no neck and truncal control and severe visual impairment. The older sister also now shows signs of dystonia with parkinsonism.

Cerebral MRI / MR-spectroscopy

Cerebral MRI was performed in all patients (Fig. 2 and 3). MRI images of patients 6 and 7 are not shown due to low quality but demonstrated hypomyelination and thin corpus callosum. The MR images of patients 1, 2, 4 and 5 show a consistent pattern with hypoplastic corpus callosum, hypomyelination, enlarged ventricles and generalized brain atrophy. Patient 3, who was only 4 months old at the time of MR imaging, showed no clear abnormalities, however, a relatively

thin corpus callosum and mildly enlarged ventricles were noted. In two patients, repeated MR imaging after 3 and 6 years demonstrated progressive cerebral and cerebellar atrophy (Fig. 2). MR-spectroscopy in patient 1 showed a markedly reduced tNAA peak (Fig. 2M) compared to the control (Fig. 2N), a sign of reduced vital neuronal tissue.

Whole exome sequencing

Whole exome sequencing in patients 1 and 2 confirmed a non consanguineous background (Suppl. tab. 1). Based on the phenotype similarity with Rett syndrome (MIM#312750), we examined the well covered *MECP2* and detected no variants. In addition, we analyzed the TOP 10 MECP2 interactors generated by the STRING version 11.0 and found no rare variant in the well covered coding sequence regions.

Extensive analysis of the dataset revealed no common rare homozygous variants, but possible compound heterozygous variants in *TAF8*, *CLCN1*, *TG* and *ZNF556,* common in both patients (Suppl. tab 1 and 2). Detected variants in the genes *CLCN1*, *TG* and *ZNF556* were known dbSNP listed polymorphisms observed with a high occurrence in the Genome Aggregation Database (gnomAD). In addition, most of the *CLCN1*, *TG* and *ZNF556* detected variants were predicted to have no functional impact demonstrated by a low median rank score above all bioinformatic predictors (Suppl. tab. 3). Both variants in the *TAF8* gene showed a high median rankscore (0.97; 0.94) based on their impact in splicing of *TAF8* predicted by ADA and RF score. *TAF8* intronic variant GRCh38:chr6:g.42050590A>G [c.45+4A>G (NM_138572.2)] was predicted to have an effect on the donor site of exon 1 and the exonic *TAF8* variant GRCh38:chr6:g.42057513G>A [c.489G>A (NM_138572.2)] was predicted to have an effect on the donor site of exon 5. They affect a highly conserved site across multiple species (GERP score) and are listed minor [rs554917914: c.489G>A (NM_138572.2)] or nonexistent [c.45+4A>G (NM_138572.2)] in the human genetic database gnomAD. So far, both *TAF8* variants are not known to be disease-causing (HGMD, ClinVar).

Sanger sequencing of the *TAF8* gene confirmed heterozygosity for the c.45+4A>G variant in the siblings and the mother and heterozygosity for the c.489G>A variant in the siblings and the father (Suppl. fig. 1A).

Exome sequencing patients 3-7 demonstrated homozygosity for the c.781–1G>A mutation in *TAF8*. No further mutations in genes associated with neurodevelopmental or neurodegenerative disorders were found. All mutations were confirmed by Sanger sequencing. Sanger sequencing demonstrated heterozygosity of the parents in family 3 for the c.781–1G>A mutation in *TAF8* (Suppl. fig. 1B, C). Sanger sequencing of family 4 is not shown.

Analysis of the effect of the novel splice site variants in patients 1 and 2

To investigate the effect of the splice mutations c.45+4A>G and c.489G>A, total RNA was isolated from fibroblasts of patient 1 and analyzed by RT-PCR with primers spanning exon 1 to the junction of exon 7 and 8. Three fragments were detected, extracted and sequenced (Suppl. fig. 2A, primer sequences: suppl. table 4). Analysis of the smallest fragment (*c*) revealed that exon 5 (125bp) and exon 6 (148bp) were deleted (p.Pro122Leufs*98) (Suppl. fig 2B). Sequencing of the middle PCR fragment (*b*) showed that exon 4 was joined to exon 6 because of skipping of exon 5. This aberrant junction resulted in a frameshift and introduced a premature stop codon (TGA), which was four codons downstream from the codon for Alanine at position 121 (p.Pro122Hisfs*4). Sequencing the larger PCR fragment (*a*), showed the wt TAF8 sequence from exon1 to exon7, including the wt c.489G. These data together indicate that the two shorter fragments correspond to transcripts from the paternal allele with the c.489G>A *TAF8* variant, whereas the largest PCR product not bearing this variant most likely originates from the maternal allele carrying the *TAF8* c.45+4A>G variant. Based on these findings, we speculated that the c.45+4A>G splice site mutation might be leaky leading to production of a normal transcript and possibly also an aberrant unstable small transcript that is undetectable under our experimental conditions. This being the case, a more than 50% reduction in normal TAF8 mRNA would be expected. To test this hypothesis, RT coupled with quantitative PCR (RT-qPCR) was performed using primers specifically designed to estimate the residual level of normally-spliced transcript derived from the c.45+4A>G variant (Fig. 4A). The analysis indeed demonstrated a 70% reduction of *TAF8* mRNA transcripts, indicating that the c.45+4A>G variant is a leaky splice site mutation that results in the production of reduced amounts of wild-type TAF8 mRNA.

To analyze the amount of endogenous TAF8 protein, whole cell extracts were prepared from control and patient fibroblasts and probed with two different polyclonal TAF8 antibodies, pAb3478 and Abcam 204894, which were generated against the entire TAF8 protein or to detect the N-terminal peptide sequence of TAF8, respectively[15]. Western blot analysis analyses using these two different antibodies showed a significant reduction in TAF8 protein in the patient's fibroblasts (Fig. 4B).

Next, we investigated the subcellular localization of TAF8 in patient 1 fibroblasts. Immunofluorescent micrographs demonstrated that the TAF8 immunoreactivity was strongly decreased in patients' fibroblasts, especially in the nuclei (Fig 4C, inset a and b).

Reduced TAF10 expression in nuclear fractions of the patient with a TAF8 variant

Studies have shown that TAF8 regulates the nuclear import of TAF10, and they assemble co-translationally[26, 27]. Given this interaction, reduced expression of TAF8 in patients may affect the level of nuclear localization of TAF10. Therefore, we sought to determine the expression of TAF8 and TAF10 in different subcellular fractions. Consistent to previous assessment in whole cell extraction (Fig. 4B), the expression of TAF8 in the patient was greatly reduced in both the cytoplasmic and nuclear fractions (Fig. 4D, E, F). Importantly, we also observed that expression of TAF10 was reduced in the nuclear fraction (NF) of the patient, while the cytoplasmic fraction (CF) remained unaffected compared to the control (Fig. 4D, G). As TAF8, TAF10 and TAF2 form a complex in the cytoplasm which is then nucleated by TAF8, we can assume that less functional TAF8-TAF10-TAF2 complex will be incorporated into the core-TFIID complex in the nucleus of the patient[28].

Discussion

In 2018 El-Saafin et al. described a patient with a homozygous splice site mutation, c.781–1G>A, in *TAF8*[15]. Analysis in fibroblasts showed that the TAF8 protein was undetectable by western blotting and mass spectrometry analyses of purified TFIID complexes, presumably due to the instability of the resulting protein. This finding was highly unexpected as germline deletion of Taf8 in the mouse has been reported to cause embryonic lethality at E4.0[29]. El-Saafin et al. then tested the effect of an inducible deletion of *Taf8* in mouse embryonic stem cells (mESCs) and also found that it leads to cell death[15]. In the same study the authors analyzed mESCs with low residual levels of TAF8 protein and found that they can survive and divide similarly to controls[15]. Therefore, they speculated that the survival of the patient might be due to very low levels of residual TAF8 protein undetectable by the antibodies used. Five of the patients presented in the present study carry the same homozygous splice site mutation, c.781–1G>A, in the *TAF8* gene. Sibling patients 1 and 2, however, bear novel compound heterozygous splice site mutations in *TAF8* (c.45+4A>G and c.489G>A). To study the pathogenicity of the two novel splice site mutations we first analyzed *TAF8* mRNA in fibroblasts from patient 1 and identified 3 transcripts. By sequencing the RT-PCR products, we were able to show that the 2 smaller fragments were a result of the c.489G>A variant while the large fragment showed a wt *TAF8* sequence. To test if the c.45+4A>G also has a pathogenic effect, we next quantified *TAF8*

mRNA and performed Western blotting using two TAF8 specific antibodies. Supporting the pathogenicity of the splice site variant we found a reduction of more than 50% for both *TAF 8* mRNA and TAF8 protein compared to wt fibroblasts. Our results therefore indicate that c.45+4A>G severely affects TAF8 protein expression. Immunohistochemistry using TAF8 specific antibodies demonstrated an overall reduction of TAF8 protein that was more pronounced in the nucleus than in the cytoplasm. This finding was confirmed by the analysis of subcellular localization of the TAF8 protein that again showed an overall reduction of TAF8 protein with a greater reduction in the nuclei than in cytoplasm. Consistent with these findings, TAF10, a subunit of TFIID that is dependent on TAF8 for transport into the nucleus, was also reduced in the nucleus, while its abundance in cytoplasm was comparable with the control fibroblasts[28]. Together, these findings indicate that the compound heterozygous mutations detected in patients 1 and 2 exert a less severe effect than that reported for the c.781–1G>A mutation for which no detectable TAF8 protein was found. Nevertheless, the clinical phenotype in all seven patients is equally severe supporting the notion that the c.781–1G>A also leads to low levels of residual TAF8 protein, that escape detection.

The seven patients with *TAF8* mutations described in this article, aged between 2 and 18 years, show a similar clinical phenotype. All were born after a normal pregnancy without dysmorphic features. The postnatal period was marked by severe feeding problems involving poor sucking and swallowing and ongoing difficulties have necessitated tube feeding. Poor growth has led to short stature and microcephaly. Motor development is severely impaired in all patients, initially by axial hypotonia in infancy and later by progressive spasticity in childhood and adolescence. None of the patients have achieved sitting or walking and head control remains poor. All patients show severe intellectual disability with no language development and poor visual fixation. Except for the youngest patient, all developed epilepsy in early childhood. Cranial MRI is characterized by hypomyelination, enlarged ventricles and a thin corpus callosum. So far, repeated MRI examinations have only been performed in the German siblings and show a clearly progressive atrophy in both, affecting grey and white matter. MR-spectroscopy performed in one of the siblings showed a reduced tNAA peak in cortical grey matter indicating neuronal loss. This indicates that the TAF8-associated disorder has a neurodegenerative disease course. Clinically this progression is characterized in the patients by progressive spasticity, reduction of spontaneous movements and development of epilepsy.

So far eight disorders associated with defects in subunits of TFIID have been described (Table 2). Two of them, X-linked dystonia-parkinsonism (MIM# 314250) and spinocerebellar ataxia 17 (MIM# 607136) have a more complex genetic background. X-linked dystonia-parkinsonism

is caused by a retrotransposon inserted in intron 32 of the *TAF1* gene and spinocerebellar ataxia 17 by expansion of a CAG repeat in the *TBP* gene. There is still a debate as to whether these disorders are caused by a gain or loss of function of TAF1 and TBP, respectively. Moreover, both disorders are clinically distinct from the other disorders that are related to defects in subunits of TFIID as they appear later in life and are associated primarily with movement disorders. Conversely, mutations that likely cause a loss of function in the TFIID subunits TAF1, TAF2, TAF6, TAF13, TBP and, as we confirm in this article, TAF8, are associated with overlapping congenital disorders characterized by microcephaly, developmental delay and intellectual disability. As only very few patients with such defects have been described, the full clinical spectrum is not yet known, but so far, the organ that is primarily affected is the central nervous system. This is an unexpected finding because TFIID is expressed in all tissues and essential for cell survival. Analysis of a zebrafish *TAF1* knockout model demonstrated more than 6000 differentially expressed genes[30]. There are, however, other examples of severe neurodevelopmental disorders that result from complex misregulated gene expression expected to affect all organs of the body. Rett Syndrome, one of the most common monogenetic causes of mental retardation in females, is caused by mutations in *MECP2*. The mutations have been shown to lead to misregulation of thousands of genes[31]. Two further examples are Angelman syndrome and Cornelia de Lange syndrome[13, 32]. Phenotypically these disorders are primarily characterized by microcephaly, brain atrophy, epilepsy and severe intellectual disability. In fact, the presenting diagnosis of patient 1 and 2 in this article was Rett syndrome and, consequently, exome sequencing was performed as part of a project aimed at detecting new Rett syndrome related genes. This emerging phenotype of disorders associated with misregulated protein expression most likely reflects a selective vulnerability of the CNS.

In two of the patients with *TAF8* variants, cerebral MRI was repeated and clearly showed progressive brain atrophy. Progressive symptoms have also been described in patients with variants in *TAF1, TAF2* and *TBP*. A detailed analysis of spatial-temporal expression of TFIID subunits might help to understand the dynamic of the disease and in the future open avenues for targeted treatment.

A limitation of our study is that the biochemical experiments were performed in fibroblasts although the main pathology is seen in the brain. To better understand the pathophysiology of this novel disorder we intend to continue our experiments in neurol progenitor cells derived from fibroblasts. Another possible limitation of our biochemical experiments is that expression of the housekeeping genes used for normalization but also *TAF8* itself might be affected by the TAF8 defect. However, this seems unlikely as in the original study on the c.781-1G > A variant

no significant changes in genome-wide Pol II occupancy and Pol II-dependent pre-mRNA transcription were seen[15].

In conclusion, we confirm that biallelic mutations in *TAF8* cause a very severe neurodevelopmental disorder characterized by almost no development, microcephaly and epilepsy. MRI/ MRS evidence in two patients indicates a neurodegenerative course. The phenotype of the *TAF8*-associated disorder has a high overlap with that seen in patients with other genetic defects caused by loss of function mutations affecting genes encoding TFIID subunits.

Acknowledgements

We would like to thank Lazlo Tora and Didier Devys for helpful discussions and critically reading the manuscript

Funding

This work was supported by the German Research Foundation (DFG, Ga354/14-1, HU 941/2-5), and the Niedersächsisches Ministerium für Wissenschaft und Kultur (11-76251-12-4/12 (ZN2938))

The family 3 (patients 4 and 5) was collected as part of the SYNaPS Study Group collaboration funded by The Wellcome Trust and strategic award (Synaptopathies) funding (WT093205 MA and WT104033AIA). This research was conducted as part of the Queen Square Genomics group at University College London, supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. SE and HH were supported by an MRC strategic award to establish an International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) MR/S005021/1'.

Competing interests

The authors report no competing interests

Supplementary material

Supplementary material is available at Brain online

References

1. Tora L. A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev*. Mar 15 2002;16(6):673-5. doi:10.1101/gad.976402
2. Koide R, Kobayashi S, Shimohata T, et al. A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum Mol Genet*. Oct 1999;8(11):2047-53. doi:10.1093/hmg/8.11.2047
3. Yang S, Li XJ, Li S. Molecular mechanisms underlying Spinocerebellar Ataxia 17 (SCA17) pathogenesis. *Rare Dis*. 2016;4(1):e1223580. doi:10.1080/21675511.2016.1223580
4. Lee JY, Cho YH, Hallford G. Delineation of subtelomeric deletion of the long arm of chromosome 6. *Ann Hum Genet*. Nov 2011;75(6):755-64. doi:10.1111/j.1469-1809.2011.00675.x
5. Lee LV, Pascasio FM, Fuentes FD, Viterbo GH. Torsion dystonia in Panay, Philippines. *Adv Neurol*. 1976;14:137-51.
6. Aneichyk T, Hendriks WT, Yadav R, et al. Dissecting the Causal Mechanism of X-Linked Dystonia-Parkinsonism by Integrating Genome and Transcriptome Assembly. *Cell*. Feb 22 2018;172(5):897-909 e21. doi:10.1016/j.cell.2018.02.011
7. Evidente VG, Advincula J, Esteban R, et al. Phenomenology of "Lubag" or X-linked dystonia-parkinsonism. *Mov Disord*. Nov 2002;17(6):1271-7. doi:10.1002/mds.10271
8. Cheng H, Capponi S, Wakeling E, et al. Missense variants in TAF1 and developmental phenotypes: challenges of determining pathogenicity. *Hum Mutat*. Oct 23 2019;doi:10.1002/humu.23936
9. Hellman-Aharony S, Smirin-Yosef P, Halevy A, et al. Microcephaly thin corpus callosum intellectual disability syndrome caused by mutated TAF2. *Pediatr Neurol*. Dec 2013;49(6):411-416 e1. doi:10.1016/j.pediatrneurol.2013.07.017
10. Najmabadi H, Hu H, Garshasbi M, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*. Sep 21 2011;478(7367):57-63. doi:10.1038/nature10423
11. Alazami AM, Patel N, Shamseldin HE, et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep*. Jan 13 2015;10(2):148-61. doi:10.1016/j.celrep.2014.12.015
12. Tuc E, Bengur FB, Aykut A, Sahin O, Alanay Y. The third family with TAF6-related phenotype: Alazami-Yuan syndrome. *Clin Genet*. May 2020;97(5):795-796. doi:10.1111/cge.13711
13. Yuan B, Pehlivan D, Karaca E, et al. Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J Clin Invest*. Feb 2015;125(2):636-51. doi:10.1172/JCI77435
14. Tawamie H, Martianov I, Wohlfahrt N, et al. Hypomorphic Pathogenic Variants in TAF13 Are Associated with Autosomal-Recessive Intellectual Disability and Microcephaly. *Am J Hum Genet*. Mar 2 2017;100(3):555-561. doi:10.1016/j.ajhg.2017.01.032
15. El-Saafin F, Curry C, Ye T, et al. Homozygous TAF8 mutation in a patient with intellectual disability results in undetectable TAF8 protein, but preserved RNA polymerase II transcription. *Hum Mol Genet*. Jun 15 2018;27(12):2171-2186. doi:10.1093/hmg/ddy126
16. Frahm J, Bruhn H, Gyngell ML, Merboldt KD, Hanicke W, Sauter R. Localized high-resolution proton NMR spectroscopy using stimulated echoes: initial applications to human brain in vivo. *Magn Reson Med*. Jan 1989;9(1):79-93. doi:10.1002/mrm.1910090110

17. Natt O, Bezkorovaynyy V, Michaelis T, Frahm J. Use of phased array coils for a determination of absolute metabolite concentrations. *Magn Reson Med*. Jan 2005;53(1):3-8. doi:10.1002/mrm.20337
18. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. Aug 15 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352
19. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. Sep 2010;20(9):1297-303. doi:10.1101/gr.107524.110
20. Rimmer A, Phan H, Mathieson I, et al. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet*. Aug 2014;46(8):912-918. doi:10.1038/ng.3036
21. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11 10 1-11 10 33. doi:10.1002/0471250953.bi1110s43
22. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. Apr-Jun 2012;6(2):80-92. doi:10.4161/fly.19695
23. Efthymiou S, Salpietro V, Malintan N, et al. Biallelic mutations in neurofascin cause neurodevelopmental impairment and peripheral demyelination. *Brain*. Oct 1 2019;142(10):2948-2964. doi:10.1093/brain/awz248
24. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. Dec 2001;25(4):402-8. doi:10.1006/meth.2001.1262
25. Beg AA, Finco TS, Nantermet PV, Baldwin AS, Jr. Tumor necrosis factor and interleukin-1 lead to phosphorylation and loss of I kappa B alpha: a mechanism for NF-kappa B activation. *Mol Cell Biol*. Jun 1993;13(6):3301-10. doi:10.1128/mcb.13.6.3301
26. Kamenova I, Mukherjee P, Conic S, et al. Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat Commun*. Apr 15 2019;10(1):1740. doi:10.1038/s41467-019-09749-y
27. Soutoglou E, Demeny MA, Scheer E, Fienga G, Sassone-Corsi P, Tora L. The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Mol Cell Biol*. May 2005;25(10):4092-104. doi:10.1128/MCB.25.10.4092-4104.2005
28. Trowitzsch S, Viola C, Scheer E, et al. Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat Commun*. Jan 14 2015;6:6011. doi:10.1038/ncomms7011
29. Voss AK, Thomas T, Petrou P, Anastassiadis K, Scholer H, Gruss P. Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. *Development*. Dec 2000;127(24):5449-61.
30. Gudmundsson S, Wilbe M, Filipek-Gorniok B, et al. TAF1, associated with intellectual disability in humans, is essential for embryogenesis and regulates neurodevelopmental processes in zebrafish. *Sci Rep*. Jul 24 2019;9(1):10730. doi:10.1038/s41598-019-46632-8
31. Della Ragione F, Filosa S, Scalabri F, D'Esposito M. MeCP2 as a genome-wide modulator: the renewal of an old story. *Front Genet*. 2012;3:181. doi:10.3389/fgene.2012.00181
32. Rotaru DC, Mientjes EJ, Elgersma Y. Angelman Syndrome: From Mouse Models to Therapy. *Neuroscience*. Oct 1 2020;445:172-189. doi:10.1016/j.neuroscience.2020.02.017

Figure legends

Figure 1. Pedigrees and gene structure. (**A**) Pedigree of the German family with patient 1 and patient 2. (**B**) Pedigree of the United States family with patient 3. (**C**) Pedigree of the Moroccan family with patient 4 and patient 5. (**D**) Pedigree of the Iranian familywith patients 6 and 7. E) Gene structure of *TAF8.* Position of splice site variants is indicated by a red arrow. Results of the sanger sequencing can be found in Supplementary Figure 1.

Figure 2. Cranial MRI/MRS images of patients 1 and 2. (**A, D, G, J**) T1 weighted sagittal images and (**B, C, E, F, H, I, K, L**) T2-weighted axial images of patients 1 and 2. (**A-C**) Images of patient 1 at age 16months demonstrating a thin corpus callosum hypomyelination and mildly enlarged ventricles. (**D-F**) Images of patients 1 at age 10 years demonstrating progressive global cerebral and cerebellar atrophy with grossly enlarged ventricles. (**G-I**) Images of patient 2 at age 21 months showing a thin corpus callosum periventricular hypomyelination and enlarged ventricles. (**J-L**) Follow-up MRI of patient 2 at age 5 years demonstrating, like her sister, clearly progressive cerebral and cerebellar atrophy. MR spectrum from cortical GM of patient 1 at age 7 years (M) and of a healthy 7-year-old female (N). The tNAA peak was distinctly decreased in the patient compared to the control and less clearly the tCr peak indicating loss of viable neuronal tissue.

Figure 3. MRI images of patients 3, 4 and 5. (**A, G**) T1 weighted images, (**B, C, F, H, I**) T2 weighted images, (**D, E**) Fluid-attenuated inversion recovery (FLAIR) images. (**A-C**) Images of patient 3 at age 4 months showing mild cerebral and cerebellar atrophy. (**D-F**) images of patient 4 at age 17 years demonstrating enlarged ventricles and thin corpus callosum. (**G-I**) images of patient 5 at age 10 months demonstrating mildly enlarged ventricles.

Figure 4. Biochemistry analysis of TAF8 patient carrying c.45+4A>G and c.489G>A variants. (**A**) *Top*: Schematic illustrating the location of primers used for the transcript analysis. Forward primer 83F locates on Exon 5 and reverse primer 84R spans between junction of exon 5 and 6. *Bottom*: qRT–PCR analysis of TAF8 mRNA expression in primary fibroblast cell line from a control and a patient. Expression is normalized to that of beta actin *ACTB*. Percentage of mRNA is equal to $2^{-\Delta\Delta CT}$ and normalized relative to control. Data are given as means ±

SD, $n = 3$ independent experiments. Statistical differences were obtained with unpaired Welch's t-test: $*P \leq 0.05$, $**P \leq 0.01$. (**B**) Western blot analysis (cropped blot images) of control and patient fibroblast (whole cell lysates) probed with the indicated antibodies for TAF8. TAF8 densitometry is shown on the right. Note no truncated proteins were detected. (**C**) Representative Immunofluorescence for TAF8 (red) in control and patient fibroblast. TAF8 was detectable in the cytoplasm but mainly localized to nuclei (DAPI, blue) in both control and patient's fibroblast. The arrows depict colocalization. Magnified view in the insets correspond to single cell outlined by white rectangle with letter a and b. Note the TAF8 immunoreactivity is strongly decreased in patient's fibroblast. Scale bars, 20μM. (**D**) Western blot analysis (cropped blot images) of TAF8 (two different antibodies as indicated) and TAF10 in the whole cell extract (WC), cytoplasmic fraction (CF) and nuclear fraction (NF). GAPDH and Histone H3 were used as cytoplasmic and nuclear marker protein expression, respectively. (**E**) Densitometry of TAF8 p3478, (**F**) TAF8 Abcam 204894 and (**G**)TAF10 in different subcellular fractions. Note that the lanes of whole cell extract in TAF8 and TAF10 blots were overexposed in order to show the relative protein level in CF and NF. See Supplementary Figure 3 for full blots and blots with shorter exposure time.