

Genome editing in induced pluripotent stem cells rescues TAF1 levels in X-linked dystonia-parkinsonism

Aleksandar Rakovic^{1*}, PhD; Aloysius Domingo^{1*}, MD, PhD; Karen Grütz^{1*}, PhD, Leonora Kulikovskaja¹, MSc; Philipp Capetian¹, MD; Sally A. Cowley², PhD; Norbert Brüggemann¹, MD; Raymond Rosales³, MD; Dominic Jamora⁴, MD; Arndt Rolfs⁵, MD; Philip Seibler¹, PhD; Ana Westenberger¹, PhD; Inke König⁶, PhD; Christine Klein¹, MD

¹Institute of Neurogenetics, University of Lübeck, Lübeck, Germany

²Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom

³Departments of Neurology and Psychiatry, University of Santo Tomas, Manila, Philippines

⁴Department of Neurosciences, College of Medicine, University of the Philippines-Philippine General Hospital, Manila, Philippines

⁵Albrecht-Kossel-Institute for Neuroregeneration (AKos), University of Rostock, Rostock, Germany

⁶Institut für Medizinische Biometrie und Statistik, University of Lübeck, Lübeck, Germany

*These authors contributed equally

Corresponding Author: Prof. Christine Klein, Institute of Neurogenetics, University of Lübeck, 23538 Lübeck, Germany , Tel: +49-451-31018200, email: christine.klein@neuro.uni-luebeck.de

Word count: 244 (abstract) 3,625 (text), Figures: 5

Running title: Excision of the SVA insertion rescues TAF1 expression

Keywords: Parkinsonism, Dystonia, induced pluripotent stem cells, genome editing

Conflict of interest: The authors declare no conflict of interest.

Abstract

Objective

The most likely genetic cause of X-linked dystonia-parkinsonism (XDP), a neurodegenerative movement disorder endemic to the Philippines, is a 2,672-bp SVA retrotransposon insertion within intron 32 of the TAF1 gene. We sought to investigate whether i) TAF1 expression is altered in induced pluripotent stem cells (iPSC) and differentiated neuronal models and ii) the excision of the SVA insertion in iPSCs restores normal TAF1 expression.

Methods

Expression of TAF1 and its neuronal isoform was determined in iPSCs and in iPSC-derived cortical neurons and spiny projection neurons (SPNs) using qPCR. CRISPR/Cas9-based genome excision of the SVA was performed on iPSCs from three XDP patients. Edited and unedited iPSCs from XDP patients and controls were differentiated into cortical neurons and SPNs and TAF1 expression compared across groups.

Results

TAF1 was reduced in patient-derived iPSCs ($p < 0.05$) and SPNs ($p < 0.01$). After genome editing, we observed higher TAF1 expression in edited in comparison to unedited iPSCs ($p < 0.0001$). In edited SPNs, TAF1 expression also increased, albeit not reaching statistical significance. No expression differences were observed in cortical neurons.

Conclusions

i) iPSCs and differentiated SPNs recapitulate TAF1 reduction in XDP, which is mostly likely due to the SVA insertion. ii) TAF1 represents a tractable molecular phenotype of XDP that can be driven with the excision of the SVA insertion. iii) Successful rescue of the molecular phenotype at the iPSC level in an endogenous, CRISPR/Cas9-corrected model serves as a proof of principle which may successfully be transferred to other inherited neurodegenerative diseases.

Introduction

X-linked dystonia-parkinsonism (XDP) is a severe, neurodegenerative disorder that is currently described exclusively in individuals of Filipino descent due to a genetic founder effect. The molecular events leading to XDP are not completely understood, in part due to complete cosegregation of all known disease-specific genetic alterations in patients, and the complete absence of these variants in controls ¹. These DNA alterations include five single nucleotide substitutions, one 48-bp deletion, and a 2,672-bp SINE-VNTR-Alu retrotransposon insertion ^{2,3}. All seven alterations lie in introns of, or in the intergenic region surrounding the *TAF1* gene.

Given the critical role of *TAF1* in RNA polymerase II-dependent transcription ⁴, previous functional studies have focused on transcriptional dysregulation as the disease mechanism in XDP. Reduced expression of *TAF1* and of a neuron-specific variant, *nTAF1*, at the mRNA level was described in the caudate of an XDP patient ²; and reduced expression of *TAF1* in blood and fibroblasts of patients ⁵, and of various *TAF1* transcripts in neural stem cells derived from induced pluripotent stem cells (iPSCs) ⁶ have also been documented. However, in all of these studies, the patient-derived biomaterials and cells carried all seven XDP-specific DNA variants in tandem, making it difficult to deconvolute independent effects and determine which specific genetic element is responsible for the observed reduction of *TAF1* or *nTAF1* at the mRNA level. Notably, the most plausible candidate mutation is the SVA retrotransposon insertion in intron 32, given its proximity to exon 34', the exon that is used by *nTAF1* ².

Genome editing using CRISPR/Cas9 technology allows sequence-specific genome manipulations in many organisms and cellular models ⁷. The method involves the introduction of targeted double-stranded breaks (DSBs) in DNA, which are typically repaired by non-homologous end-joining (NHEJ) and results in insertions, deletions or nucleotide substitutions. Using a dual-guide RNA (gRNA) approach, where Cas9 nuclease is coexpressed with two gRNAs targeting two ends of a genomic region of interest, it is possible to generate deletions of segments of DNA in CRISPR-edited cells that differ from their parent line with respect to the excised region, but otherwise retain the full genetic background of the donor ⁸.

In order to investigate whether the 2,672-bp SVA retrotransposon insertion is causal to the reduction of *TAF1* expression in XDP, we utilized CRISPR/Cas9-based genome editing to generate XDP cell lines with the SVA retrotransposon insertion excised. We performed the editing in iPSCs, which allows for later neural induction ⁹, and further differentiation into distinct regional subtypes such as cortical ¹⁰ and spiny projection neurons (SPNs) ¹¹.

Materials and Methods

iPSC lines, neuronal cultures

All individuals gave informed consent and the Ethics Committee of the University of Lübeck approved the study. Forearm skin fibroblasts from XDP patients and controls (passage number <3) were reprogrammed into iPSCs using the CytoTune™-iPS and CytoTune®-iPS 2.0 Sendai Reprogramming kits (Invitrogen) according to the manufacturer's protocols. Generated iPSC lines were characterized by immunofluorescent staining as well as expression analysis of pluripotency markers. Genomic integrity in comparison to the parental fibroblast line was evaluated by SNP analysis on an Infinium OmniExpress-24 beadchip (Illumina).

iPSCs (passage number <5) were differentiated into cortical neurons as previously published¹². In brief, iPSCs were plated as single cells. At 95% confluency, differentiation was initiated in KSR-medium (KnockOut™ DMEM/F-12 (1X) supplemented with KnockOut™ Serum Replacement, L-Glutamine, MEM NEAA, and 2-Mercaptoethanol) supplemented with Dorsomorphine (1 µM), SB 431542 (10 µM) and Y-27632 (10 µM). Until day twelve of differentiation, the medium composition was gradually shifted from KSR-medium to neural maintenance medium (NMM; 1:1 Neurobasal® Medium (1X) and KnockOut™ DMEM/F-12 (1X), with N2-Supplement, B27-Supplement, L-Glutamine, MEM NEAA, 2-Mercaptoethanol, and insulin) supplemented with Dorsomorphine, SB 431542 and Y-27632. From day 13 to 17, the cells were cultured in NMM medium containing FGF (FGF2, basic fibroblast growth factor; 20 ng/ml) and BDNF (Brain-derived neurotrophic factor; 20 ng/ml). Neural rosettes were manually replated on day 18 in NMM with BDNF (20 ng/ml), GDNF (Glial cell-derived neurotrophic factor; 20 ng/ml), and AA (ascorbic acid; 0.2 mM). The medium was exchanged every second day until day 27 including an additional manual rosette replating on day 23. On day 28, the rosettes were dissociated using accutase and plated at desired densities in NMM (with BDNF, GDNF, and AA). Until day 43, the medium (NMM with BDNF, GDNF, and AA) was changed every three days. For final differentiation the cells were cultured in NMM.

Differentiation of SPNs was adapted from a previously published method¹¹. In brief, iPSCs were detached, resuspended in KSR-medium to form embryoid bodies in low-attachment plates.

Differentiation was induced with LDN (100 nM), SB (10 μ M), Y (10 μ M), IWP2 (1 μ M) on day 0. On day 2 the medium was replaced with a 1:1 mixture of KSR-medium and N2-medium (KnockOut™ DMEM/F-12 (1X), N2-Supplement (100X), L-Glutamine, and MEM NEAA) supplemented with LDN (100 nM), SB (10 μ M), IWP2 (1 μ M). Every other day, the medium was replaced by fresh N2-medium, supplemented with LDN (100 nM), SB (10 μ M), IWP2 (1 μ M), Purmorphamine (0.2 μ M) (day 4) and IWP2 (1 μ M), Purmorphamine (0.2 μ M) (day 6 and day 8) and non-supplemented on day 10 and day 12. On day 13, the EBs were replated on Matrigel-coated plates in N2B27 medium (1:1 Neurobasal® Medium (1X) and KnockOut™ DMEM/F-12 (1X), with N2-Supplement, B27-Supplement without vitamin A, L-Glutamine, and MEM NEAA) supplemented with BDNF (20ng/ml), GDNF (10ng/ml), cAMP (50 μ M). After day 13, the medium was changed every three days. Cells were replated manually when the outgrowing cells became too confluent on fresh Matrigel-coated plates (approximately day 30) and on Poly-D-Lysine-Laminin-coated plates for final differentiation (approximately day 45). Immunofluorescent staining and markers of mature cortical neurons and SPN were obtained from terminally differentiated cultures. RNA for expression analysis was collected using standard kits (Qiagen).

Genome editing

A dual-guide RNA strategy was used to excise the SVA retrotransposon insertion. Two gRNAs were designed, each targeting genomic sequences in the flanking regions of the SVA retrotransposon insertion (GenBank: AB191243.1, ²). gRNA1 targeted the sequence 5'-tgctgtgagcaatcatagac-3'; the PAM sequence immediately after (ggg) was 107-bp proximal to the first base of the SVA insertion. gRNA2 targeted a genomic segment distal to the insertion, 5'-tggcattgtaataaggcttg-3', with the PAM sequence (agg) was 48-bp distal to the last base of the SVA insertion. Each gRNA was cloned into the pLKO1-Puro vector (Addgene plasmid #8453). Plasmids expressing *Cas9* cDNA and the resistance gene to the antibiotic blasticidin (pCas9-Blast) were also produced. Off-target sites were predicted using the online tool Cas-OFFfinder (<http://www.rgenome.net/cas-offfinder/>). All iPSCs used for genome editing had passage numbers lower than three.

Prior to transfection, iPSCs were propagated on Matrigel-coated plates with mTESR1 feeder-free medium (Stem Cell). iPSCs were harvested using accutase treatment (Life Technologies) and passed through a 40 µm cell strainer (Falcon) to achieve single-cell status, then transfected with pCas9-Blast and both gRNA constructs, using Nucleofector kit 1 solution (Lonza) and an Amaxa Nucleofection II device (program A-023). Transfected cells were plated on Matrigel-coated wells with mTESR1 medium supplemented with Y-27632 (Santa Cruz Biotech). After 24h, selection antibiotics (blasticidin and puromycin) were introduced and maintained for 48h. Selected cells were then grown subsequently on Matrigel-coated plates until colony formation was apparent. The presence of the SVA insertion in DNA from up to 12 colonies was determined using PCR verification. Sequences of the expected junction regions were determined via Sanger sequencing. To test for off-target mutations, DNA was extracted using a standard kit (Qiagen) and predicted sites were sequenced. Primers and PCR conditions are available upon request.

Expression analysis

Total RNA was extracted from cell pellets using standard kits (Qiagen). cDNA was synthesized using the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Quantitative PCR (qPCR) was performed on a Light Cycler 480 system using the Maxima SYBR Green/Fluorescein qPCR Master Mix (Thermo Fisher Scientific). Total *TAF1* mRNA expression was measured using primers TAF1 F 5'-AGAGTCGGGAGAGCTTTCTG-3' and TAF1 R 5'-CACAATCTCCTGGGCAGTCT-3' targeting exon 32 and exon 33, respectively. To quantify the expression of *nTAF1*, a transcript-specific qPCR primers were designed. Transcript-specific forward primers were designed by using the last 16 bases of exon 34 of *TAF1*, a mismatched base at the third-to-the-last position of the primer, and a transcript-specific base at the last position, i.e. the first base of exon 34' (for the *nTAF1*-specific primer) (Fig. 3A).

Comparisons and Statistical analyses

To compare expression of *TAF1* and *nTAF1* between patient-derived and control-derived cells prior to and upon editing, data from the two runs were treated as independent observations. Expression in

iPSCs (total *TAF1* only) and cortical neurons as well as in SPNs (total *TAF1* and *nTAF1*) were normalized using either *FLAD1* or *FPGS* with *ACTB* combined as reference genes as described in ¹³ and compared using Mann-Whitney U-tests.

Western blot analysis

Proteins were extracted using RIPA buffer (ThermoFisher) and their concentrations were determined using Dc Protein Assay Kit (Biorad). Proteins were separated by SDS PAGE using NuPAGE 4–12% Bis-Tris gels (ThermoFisher). After electrophoresis, proteins were transferred to a nitrocellulose membrane (Protran) and probed with the antibodies against TAF1 (1:1000) and GAPDH (Cell Signaling Technology, 1:50000). For densitometric analyses, TotalLab software (Nonlinear Dynamics) was used.

Results

Directed cortical and striatal differentiation from iPSCs

To investigate mRNA expression levels of *TAF1*, we used previously generated induced pluripotent stem cells (iPSCs) from four healthy controls of Filipino origin and from four XDP patients. All controls and patients were males and age-matched (42 ± 5 years at biopsy taking). All patients had comparable disease age of onset (38 ± 5 years). Generated iPSCs were characterized by expression analysis as well as immunofluorescent staining of pluripotency markers (Fig. 1A) and trilineage potential (Fig. 1B). Representative immunostaining in an iPSC colony obtained from an XDP patients is shown (Fig. 1A) as well as a representative trilineage potential analysis of one control and one XDP patient-derived iPSC colony (Fig. 1B). Genomic integrity was evaluated by SNP analysis on an Infinium OmniExpress-24 beadchip (Illumina). No difference in levels of the pluripotency markers or changes in genomic integrity were observed between XDP patients and healthy controls.

All iPSC lines were differentiated in either cortical marker-enriched neuronal cultures (cortical neurons) or spiny projection neurons (SPNs) as previously published ^{11, 12}. Immunostaining after 90 days of cortical differentiation showed that the majority of cells (up to 95%) were positive for the neuronal markers MAP2 and TUJ1. In addition, markers of different cortical layers, i.e. CTIP2, TBR1, BRN2 were present in approximately 20-30% of the cells each. Immunostaining of neuronal markers and, more specifically, cortical markers in one XDP patient-derived neuronal culture is shown exemplarily (Fig. 2A).

After 20 days of patterning, striatal cultures expressed the transcription factors MEIS2 and FOXP1 (both associated with the primordial region of the striatum and precursors of spiny projection neurons, Fig. 2B, left panel). After 90 days of differentiation, the vast majority (~ 80%) of all cells were MAP2-positive neurons, of which roughly 20-30% were stained with the SPN marker DARPP-32. Between 30 and 40% of DARPP32-positive cells carried CTIP2 positive nuclei, identifying them as terminally differentiated SPNs. Immunostaining of SPN-specific markers was shown in one XDP patient-derived neuronal culture exemplarily (Fig. 2B, right panel).

TAF1 Expression analysis in iPSCs, cortical neurons, and SPNs

Previous studies demonstrated the existence of several isoforms of *TAF1*. Of particular interest for the pathology of XDP was the presence of a neuron-specific isoform, i.e. *nTAF1* which was reduced in the caudate of XDP patients². Therefore, we sought to determine relative mRNA expression levels of the total *TAF1* and *nTAF1* isoform in iPSCs, cortical neurons and SPNs between controls and XDP patients using two sets of primers, each specific for either total *TAF1* mRNA or for the *nTAF1* isoform only (Fig. 3A). We confirmed specificity of the *nTAF1*-specific set of primers using cDNA from iPSCs, cortical neurons and SPNs obtained from two healthy controls. Here, we detected a signal only in the neuronal cultures but not in iPSCs (Fig. 3B) confirming the specificity of our transcript-specific qPCR. Next, we compared total *TAF1* expression in iPSCs between the XDP group (n=4) and healthy controls of Filipino origin (n=4). Here, we observed reduction in total *TAF1* expression in the XDP group in comparison to controls ($p<0.05$, fold change=1.75) (Fig. 3B). iPSCs from three different controls and from three different XDP patients were then differentiated into cortical neurons and SPNs in two independent differentiations. Each differentiation was considered as an independent data point. In cortical neurons, no significant differences in levels of either total *TAF1* or *nTAF1* were observed between the two groups. More importantly, in SPNs, levels of *TAF1* were reduced in the XDP group in comparison to controls ($p<0.05$, fold change=1.23) (Fig. 3B). On the other hand, levels of *nTAF1* were comparable between the two groups.

Genome editing and characterization of isogenic iPSC lines

To evaluate the relevance of the retrotransposon insertion in intron 32 of the *TAF1* gene (SVA) on reduced expression of *TAF1*, we used CRISPR/Cas9-mediated gene editing to excise the SVA from control- and from XDP patient-derived iPSCs (Fig. 4A). In successfully edited XDP lines (edited XDP), the entire genomic region encompassing the SVA insertion was edited out, along with a 165 or 166-bp genomic segment immediately proximal/distal to the locus of the insertion (i.e., 120 bp proximal to the first base of the SVA insertion and 45/46 bp after the last base), representing an NHEJ-mediated deletion or 'scar'. In controls, successful editing resulted in removal of a 165bp/166bp-long genomic

'scar (edited control) (Fig. 4A). The efficiency of genome editing in control-derived iPSCs was ~30% of screened colonies and was confirmed by Sanger sequencing. PCR-based and sequencing analysis revealed successful excision of the SVA in three independent experiments using iPSC lines from three different XDP patients, in about 10-30% of colonies screened (Fig. 4B). The control- and XDP patient-derived iPSC colonies that were transfected with Cas9 and both gRNAs but failed to be edited (unedited control and unedited XDP) were also retained for further experiments. Cas-OFFinder revealed four possible off-target sites, all of which were targeted by only one gRNA. Sequencing these sites revealed only 1 of 18 randomly selected control and patient-derived iPSC colonies with a heterozygous intronic 7-bp deletion in the HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2 (HECW2) gene (Fig. 4C). No DNA segments were targeted by both gRNAs genome-wide, aside from the locus of interest on the X-chromosome. Quantification of pluripotency markers revealed their expression to be comparable to those in the parental iPSC lines (Fig. 4D). For comparison, we used fibroblasts from a healthy control that were previously reprogrammed into control iPSCs used in this study. SNP analysis on an Infinium OmniExpress-24 beadchip (Illumina) did not reveal any genetic changes which would require exclusion of the edited lines from further experiments.

Expression analysis in CRISPR-edited lines

Finally, we investigated alterations in the expression of total *TAF1* and *nTAF1* upon excision of the SVA in iPSC, cortical neurons and SPNs from both controls and XDP patients. Expression levels were measured using qPCR and normalized to combined levels of the two house-keeping genes, *ACTB* and *FLAD1*. First, we compared total *TAF1* expression either between unedited (n=5) and edited (n=3) control iPSCs or between unedited (n=12) and edited (n=10) generated from three different XDP patient-derived iPSCs. In control iPSCs, excision of the sequence proximal/distal to the SVA insertion ('scar') had no significant impact on total *TAF1* expression (Fig. 5A). On the other hand, in XDP patient-derived iPSCs we observed a significant increase in total *TAF1* expression upon ablation of the SVA (Fig. 5A). Next, two edited and four unedited colonies of a control and three edited and three unedited colonies from two XDP patient-derived iPSCs were differentiated into cortical neurons. The SPNs were

differentiated from one edited and three unedited control iPSCs and from two edited and one unedited colony from an XDP patient. For cortical neurons we performed two and for SPNs three independent differentiations. In cortical neurons, immunostaining showed that up to 95% of cells were positive for the neuronal markers MAP2 and TUJ1. In addition, markers of different cortical layers, i.e. CTIP2, TBR1, BRN2 were present in approximately 30% of the cells each. In SPNs, ~85% of all cells were MAP-2 positive, of which ~30% were labeled with the SPN marker *DARPP-32*. In control cortical neurons we observed no difference between unedited and edited neurons in levels of either total *TAF1* or *nTAF1* (Fig. 5B, upper panels). Similarly, no difference in either total *TAF1* or *nTAF1* expression was observed in XDP patients-derived cortical neurons upon excision of the SVA (Fig. 5B, upper panels). In SPNs, genome editing did not alter expressions of total *TAF1* or *nTAF1* in either the control or the XDP patient group (Fig. 5B, lower panels). Finally, we analyzed levels of the TAF1 protein in the parental, unedited, and edited iPSC-derived SPNs from a control and an XDP patient by western blotting using an antibody against TAF1 (kindly provided by Marc Timmers). In keeping with our *TAF1* mRNA expression data, we observed no differences in levels of TAF1 between parental, unedited, and edited neurons from either control- or XDP patient-derived SPNs (Fig. 5C and D). Taken together, our data indicate that excision of the SVA retrotransposon insertion in intron 32 has an impact on *TAF1* mRNA expression in iPSCs, but not in differentiated cortical neurons or SPNs.

Discussion

It currently remains a matter of debate which of the seven known genetic alterations on Xq13.1 is the disease-causing mutation in XDP. The variants occur in non-protein-coding regions of the genome, making the functional consequence of each variant difficult to predict and evaluate. Furthermore, although *TAF1* reduction has been a consistently described molecular phenotype of the disease, modelling the disease has been difficult and ascribing the observed reduction in *TAF1* to a particular DNA alteration has been impossible in previous studies, given that the genetic alterations specific to the disease occur in complete linkage disequilibrium.

In the present study, we focused on the role of the SVA retrotransposon insertion on *TAF1* expression in iPSCs and neuronal models of XDP. To isolate the contribution of this genetic element, we generated iPSC lines that were isogenic to the parental lines aside from the variant of interest, i.e. the SVA insertion. We demonstrated the robustness of this experiment, which we performed on three independent XDP donor lines, and on one control line. Our data showed an increase in *TAF1* mRNA expression in edited XDP patient-derived iPSCs in comparison to unedited counterparts, representing a rescue of *TAF1* expression as a result of the genome editing we performed, and implying the causative nature of the SVA retrotransposon insertion to reduced *TAF1* expression in XDP iPSCs. On the other hand, ablation of the SVA had no impact on expression of either total *TAF1* or its neuron-specific isoform, *nTAF1* in iPSC-derived mature neurons (whether directed towards a cortical lineage or to an SPN lineage). This apparent discrepancy in the SVA excision-related alterations in *TAF1* expression between iPSCs and iPSC-derived neurons in a primarily neurologic disease could be explained by at least two possible phenomena, i.e. i) the well-described variability upon differentiation of iPSCs into mature neurons¹⁴ could be diminishing the effects of SVA ablation, especially given that the reduction of *TAF1* in XDP and the corresponding rescue are modest (1.75 fold change in iPSCs, 1.23 fold change SPNs; rescue 1.48 fold change iPSCs in this study, 1.3 fold change in⁵). Alternatively, ii) the SVA insertion-related influence on *TAF1* expression could be a characteristic of progenitor cells (such as iPSCs), that becomes less relevant in post-mitotic cells such as mature neurons. The latter

speculation is especially intriguing, given that *TAF1* reduction was recapitulated in both iPSCs and SPNs but statistically significant rescue was only evident in iPSCs in this study. In XDP patient-derived SPNs, genome editing also resulted in increased *TAF1* expression, however not reaching significance.

Our experiments underscore the causative nature of the SVA retrotransposon insertion to reduced *TAF1* expression in XDP. *TAF1* reduction has been observed previously in XDP brains ², in peripheral cells ⁵, and in iPSC-derived models ⁶, and is now further recapitulated in our iPSCs and iPSC-derived striatal neuronal models.

Expanding on this molecular phenotype, we now demonstrate that *TAF1* expression in XDP is tractable and can be driven with the excision of one of the genetic alterations associated with the disease. The role of the SVA retrotransposon insertion to *TAF1* expression may be related to regulatory elements within or around the sequence that alter *TAF1* expression. In this regard, the sequence of the insertion is known to contain multiple CpG islands, which, when hypermethylated, could act as repressors of gene expression nearby ^{2,8}. Others have demonstrated pathogenic exon trapping by SVA elements as possible mechanisms of SVA-related alteration of host gene expression ¹⁵, although there is no evidence that this occurs in XDP.

Notably, our approach resulted in the removal of not only the SVA retrotransposon insertion but also of a small region of DNA flanking the genetic element. A precise cut that includes only the SVA insertion could not be performed due to limitations related to the absence of viable PAM sites which are necessary for Cas9 recognition and targeting. Excision of this intronic region in a control iPSC did, however, not result in a significant increase in *TAF1* mRNA expression. Nevertheless, we cannot exclude the possibility of a cryptic *TAF1*-regulatory element that is contained not only within the SVA retrotransposon insertion, but also including the flanking DNA regions.

In conclusion, our experiments show that i) iPSCs and differentiated SPNs recapitulate *TAF1* reduction in XDP ii) the SVA retrotransposon insertion in intron 32 is causal to this reduction in *TAF1*, and iii) expression can be rescued with the excision of this genetic element. Due to the proof-of-principle

nature of our study, the approach can be transferred to other monogenic diseases and thus has implications beyond XDP.

Acknowledgments

C.K. is the recipient of a career development award from the Hermann and Lilly Schilling Foundation. The project was supported by the DFG Research Unit FOR2488 (P3 to A.R., P5 to A.W.) and the Fritz Thyssen Foundation (to A.W.). We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of the SNP genotyping data. We would like to thank Marc Timmers and the CCXDP for providing us with the TAF1 antibody.

Authors' roles

A.R., A.D., and C.K. designed the study; A.R., A.D., K.G., L.K., P.C., S.A.C., A.W., A.W., A.R., R.R., D.J., and P.S. contributed to data collection and analysis. A.R., A.D., K.G., L.K., P.C., I.K., and C.K. wrote the manuscript.

Financial Disclosures of all authors (for the preceding 12 months)

Aleksandar Rakovic

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Aloysius Domingo

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Karen Grütz

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None

Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Leonora Kulikovskaja

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Philipp Capetian

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None

Expert Testimony	None
Employment	University of Lübeck
Contracts	None
Royalties	None
Other	None

Sally A. Cowley

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Oxford
Contracts	None
Royalties	None
Other	None

Norbert Bruggemann

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	Grünenthal
Grants	Else Kroener Fresenius Foundation, German Research Foundation, Collaborative Center for XDP
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None

Other	None
-------	------

Raymond Rosales

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Santo Tomas, Manila
Contracts	None
Royalties	None
Other	None

Dominic Jamora

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of the Philippines-Philippine General Hospital, Manila
Contracts	None
Royalties	None
Other	None

Arndt Rolfs

Stock Ownership in medically-related fields	Centogene AG, Rostock, shareholder and founder
---	--

Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	BMBF; European Community; intramural funds from the University of Rostock; Shire GmbH;
Intellectual Property Rights	several patents not related to the published topic
Expert Testimony	None
Employment	University of Rostock, Centogene
Contracts	None
Royalties	None
Other	None

Philip Seibler

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Ana Westenberger

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None

Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Inke König

Stock Ownership in medically-related fields	None
Consultancies	None
Advisory Boards	None
Partnerships	None
Honoraria	None
Grants	None
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	None
Other	None

Christine Klein

Stock Ownership in medically-related fields	None
Consultancies	Medical advisor to Centogene and Biogen
Advisory Boards	None
Partnerships	None
Honoraria	Wellcome Trust Review Board Scientific Advisory Board of the Else Kroener Fresenius Foundation
Grants	Movement Disorder Society, Hermann and Lilly Schilling Foundation; German Research Foundation; the BMBF; European Community; intramural funds from the University of Luebeck
Intellectual Property Rights	None
Expert Testimony	None
Employment	University of Luebeck
Contracts	None
Royalties	Oxford University Press
Other	None

References

1. Domingo A, Westenberger A, Lee LV, et al. New insights into the genetics of X-linked dystonia-parkinsonism (XDP, DYT3). *Eur J Hum Genet* 2015;23(10):1334-1340.
2. Makino S, Kaji R, Ando S, et al. Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. *Am J Hum Genet* 2007;80(3):393-406.
3. Nolte D, Niemann S, Muller U. Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. *Proc Natl Acad Sci U S A* 2003;100(18):10347-10352.
4. Thomas MC, Chiang CM. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* 2006;41(3):105-178.
5. Domingo A, Amar D, Grutz K, et al. Evidence of TAF1 dysfunction in peripheral models of X-linked dystonia-parkinsonism. *Cell Mol Life Sci* 2016;73(16):3205-3215.
6. Ito N, Hendriks WT, Dhakal J, et al. Decreased N-TAF1 expression in X-linked dystonia-parkinsonism patient-specific neural stem cells. *Dis Model Mech* 2016;9(4):451-462.
7. Paquet D, Kwart D, Chen A, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 2016;533(7601):125-129.
8. Zhao Y, Zhang C, Liu W, et al. An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci Rep* 2016;6:23890.
9. Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 2009;27(3):275-280.
10. Qi Y, Zhang XJ, Renier N, et al. Combined small-molecule inhibition accelerates the derivation of functional cortical neurons from human pluripotent stem cells. *Nat Biotechnol* 2017;35(2):154-163.
11. Stanslowsky N, Reinhardt P, Glass H, et al. Neuronal Dysfunction in iPSC-Derived Medium Spiny Neurons from Chorea-Acanthocytosis Patients Is Reversed by Src Kinase Inhibition and F-Actin Stabilization. *J Neurosci* 2016;36(47):12027-12043.

12. Grutz K, Seibler P, Weissbach A, et al. Faithful SGCE imprinting in iPSC-derived cortical neurons: an endogenous cellular model of myoclonus-dystonia. *Sci Rep* 2017;7:41156.
13. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3(7):RESEARCH0034.
14. Liang G, Zhang Y. Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* 2013;13(2):149-159.
15. Taniguchi-Ikeda M, Kobayashi K, Kanagawa M, et al. Pathogenic exon-trapping by SVA retrotransposon and rescue in Fukuyama muscular dystrophy. *Nature* 2011;478(7367):127-131.

Figures

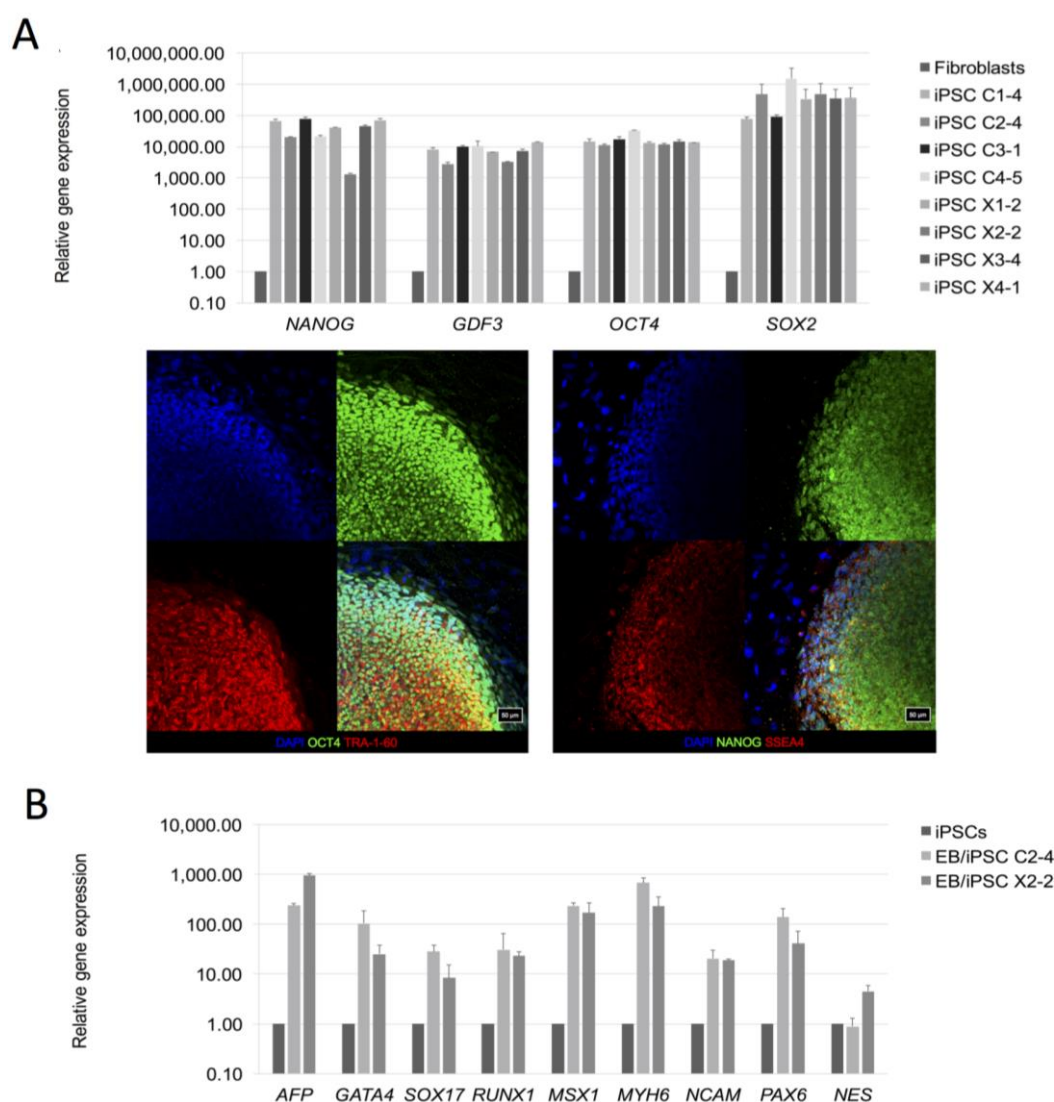
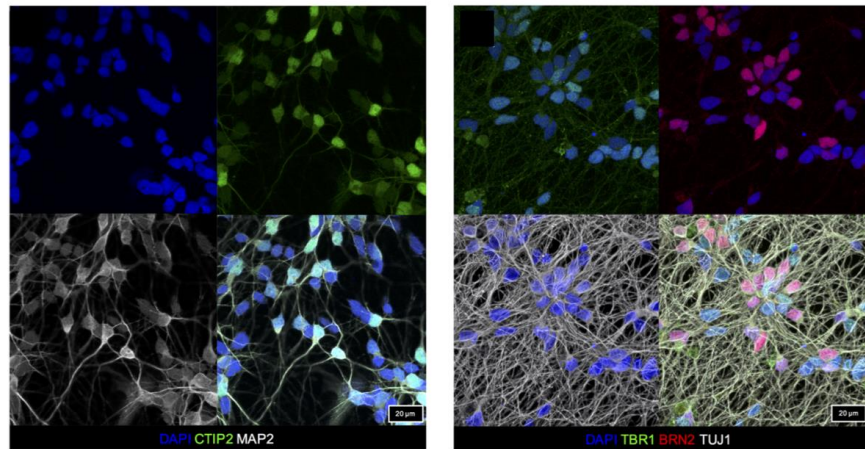


Figure 1. Characterization of iPSC colonies. (A) Relative gene expression of *NANOG*, *GDF3*, *OCT4*, and *SOX2* (pluripotency markers) in iPSCs in comparison to fibroblasts. *ACTB* was used as housekeeping gene and all expression data were normalized to fibroblasts. (C: Control, X: XDP). Representative immunostaining of expressed pluripotency markers *OCT4*, *TRA-1-60*, *NANOG*, and *SSEA4* in proliferating iPSC colonies from one healthy control. Nuclei were stained using DAPI. (B) Relative gene expression of all three germ layer markers *AFP*, *GATA4* and *SOX17* (endoderm), *RUNX1*, *MSX1* and *MYH6* (mesoderm) as well as *NCAM*, *PAX6* and *NES* (ectoderm) in spontaneously differentiated embryoid bodies in comparison to the respective iPSC in one control and one XDP patient-derived iPSC line is shown exemplarily. *ACTB* was used as housekeeping gene. (C: Control, X: XDP).

A

Cortical neurons



B

SPNs

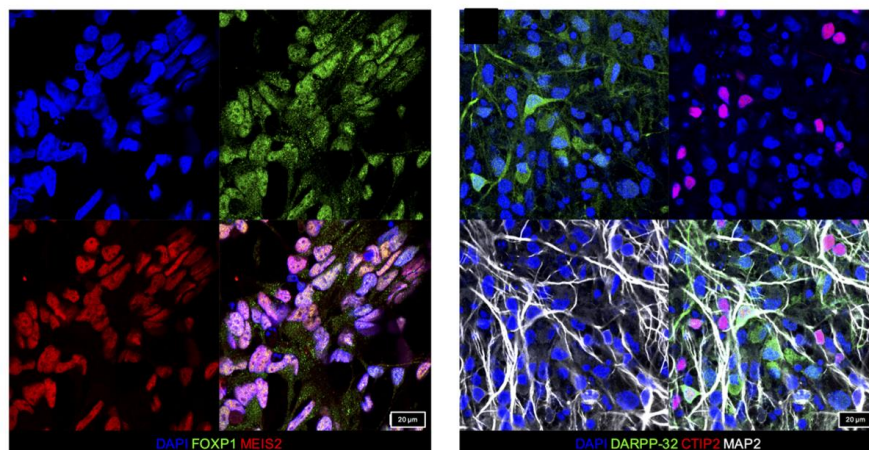


Figure 2. Characterization iPSC-derived neuronal cultures. (A) Cortical neurons after 90 days of differentiation were stained using antibodies against neuronal markers MAP2 and TUJ1, as well as CTIP2, TBR1, and BRN2, each representative of different cortical layers (V, VI, II-IV respectively). (B) Spiny projection neurons (SPNs) express transcription factors associated with the lateral ganglionic eminence (FOXP1 green, MEIS2 red). After 90 days of differentiation, MAP2-positive neurons (white) co-express DARPP-32 (green) and CTIP2 (red), a combination of markers characteristic for SPNs.

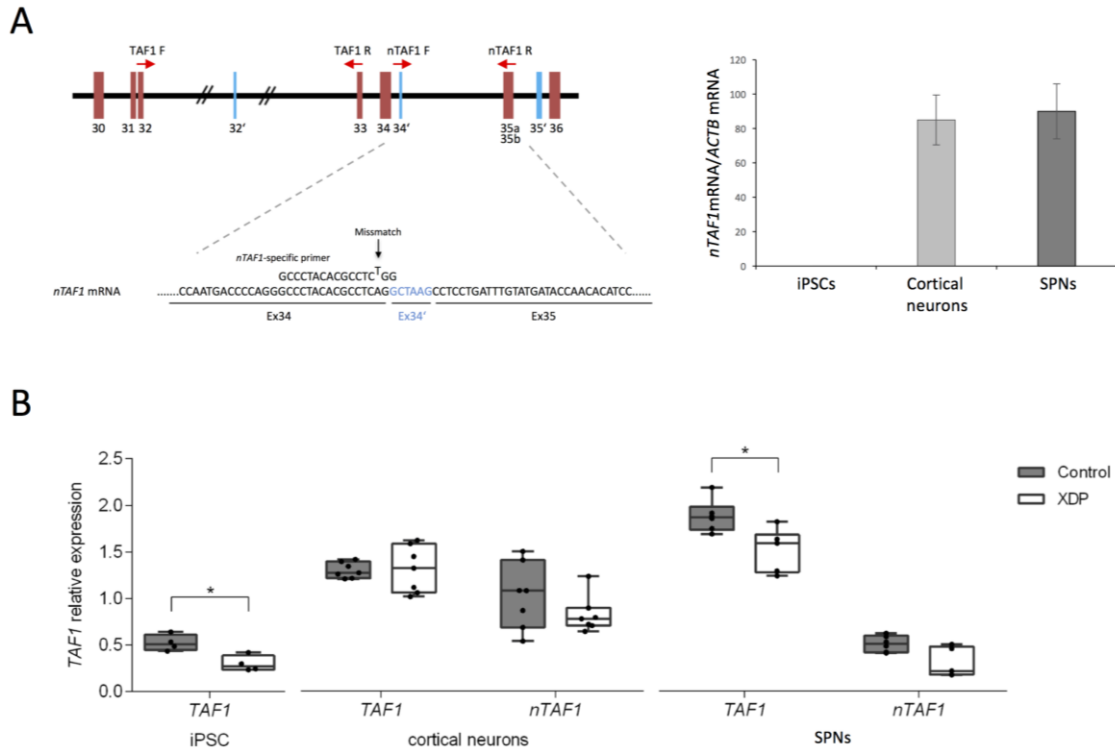


Figure 3. Reduced *TAF1* expression in XDP patient-derived iPSCs and spiny projection neurons. (A) Schematic representation of primers used to detect total *TAF1* and *nTAF1* mRNA isoforms and test of *nTAF1* mRNA-specific primers. (B) qPCR analysis showed reduced total *TAF1* expression in XDP patient-derived (white boxes) iPSC (Mann-Whitney U-test: $p < 0.05$) and SPN (Mann-Whitney U-test: $p < 0.01$) in comparison to controls (gray boxes). Furthermore, no differential total *TAF1* expression between controls and XDP patients was observed in cortical neurons. No significant differences in *nTAF1* expression were detected between controls and XDP patients. *TAF1* mRNA levels were normalized to levels of combined *ACTB* and *FPGS* mRNAs.

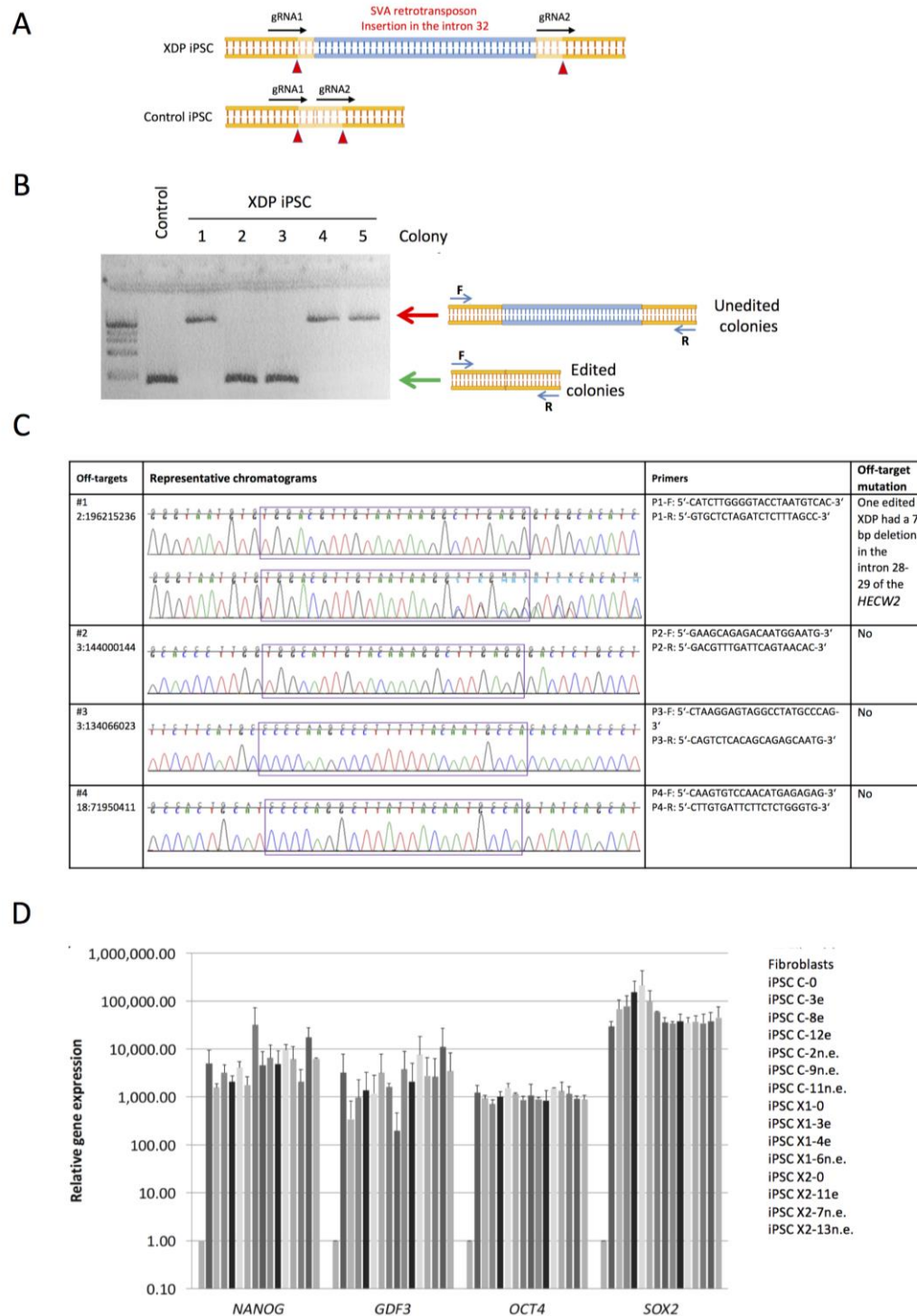


Figure 4. CRISPR/Cas9-based excision of the SVA insertion and characterization of edited control and XDP patients-derived iPSC. (A) Position and orientation of gRNAs used for excision of the SVA retrotransposon insertion in the intron 32 of the *TAF1* gene. Red arrowheads show exact positions of the cuts. Pale areas represent the ‘scar’ (B) PCR-based genotyping of iPSC colonies upon CRISPR/Cas9-based genome editing of a representative XDP patient-derived iPSC line. (C) Off-target InDels identified

in CRISPRed iPSC lines. (D) Relative gene expression of the pluripotency markers *NANOG*, *GDF3*, *OCT4*, and *SOX2* in genome-edited iPSCs in comparison to healthy control-derived fibroblasts (C: Control, X: XDP, 0: parental, e: edited, n.e: non-edited). *ACTB* served as housekeeping gene. Expression levels were normalized to fibroblasts. Error bars show standard deviation.

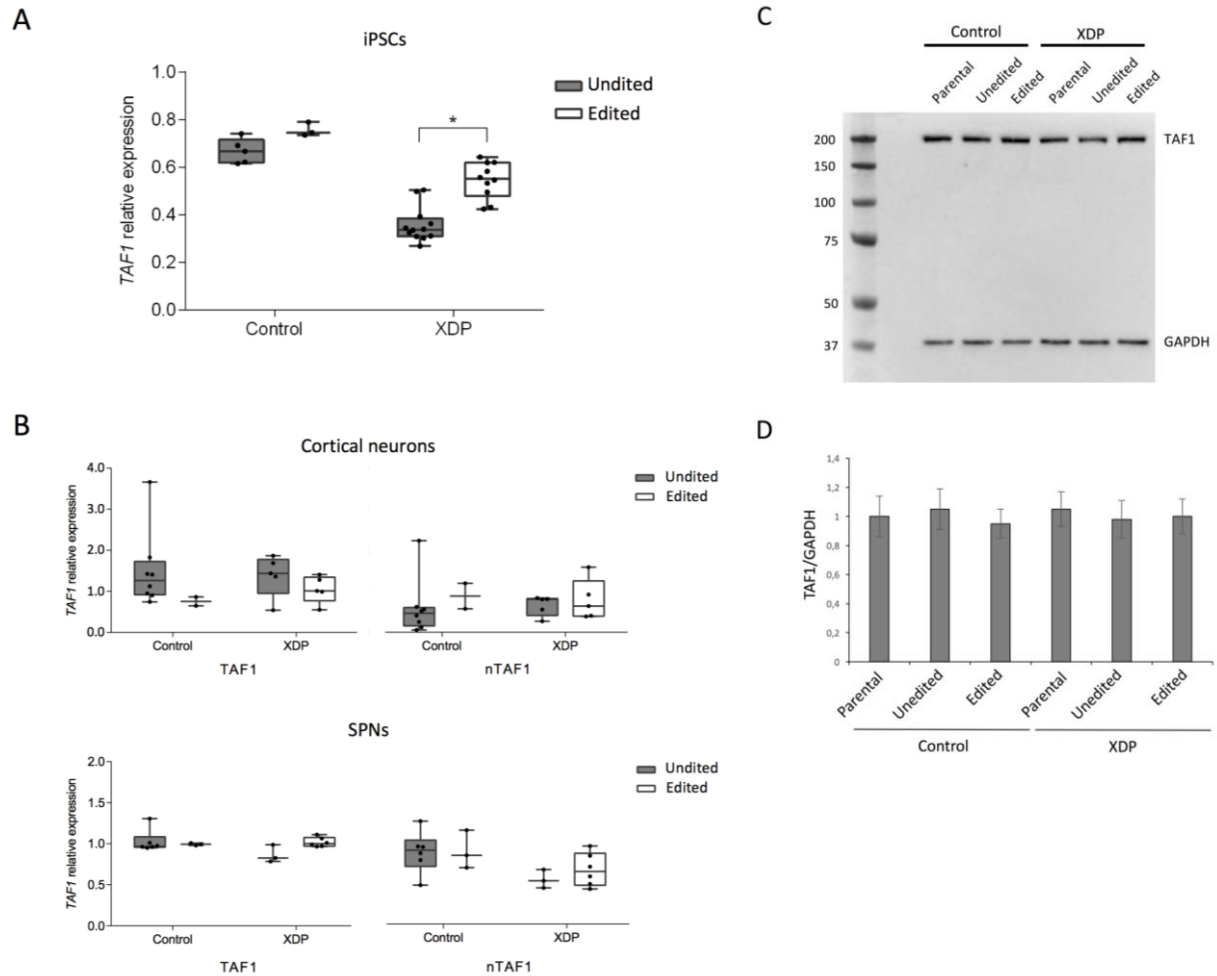


Figure 5. Excision of the SVA insertion in XDP patient-derived iPSCs increases *TAF1* expression.

CRISPR/Cas9-based excision of the SVA insertion induces increase of *TAF1* expression in the edited XDP patient-derived iPSCs (white box) when compared to unedited XDP patient-derived iPSCs (gray box) (Mann-Whitney U-test: $p < 0.0001$). In the control iPSC line, *TAF1* expression was not significantly altered upon excision of the SVA. *TAF1* mRNA levels were normalized to combined *ACTB* and *FPGS* mRNAs levels. (B) *TAF1* and *nTAF1* expression in edited and unedited SPNs derived from controls and XDP patients. (C) Western blot analysis of proteins extracted from parental, non-edited, and edited control- and XDP patient-derived SPNs. The membrane was probed with an anti-TAF1 antibody. GAPDH served as a loading control. (D) Densitometric analysis of TAF1 protein levels. Protein levels of TAF1 were normalized to the levels of GAPDH.