Disorders of the Nervous System

Pathogenic Mutation of TDP-43 Impairs RNA Processing in a Cell Type-Specific Manner: Implications for the Pathogenesis of ALS/FTLD

Kent Imaizumi, Hirosato Ideno, Tsukika Sato, Satoru Morimoto, and Hideyuki Okano

https://doi.org/10.1523/ENEURO.0061-22.2022

Department of Physiology, Keio University School of Medicine, Shinjuku, Tokyo 160-8582, Japan

Abstract

Transactivating response element DNA-binding protein of 43 kDa (TDP-43), which is encoded by the *TARDBP* gene, is an RNA-binding protein with fundamental RNA processing activities, and its loss-of-function (LOF) has a central role in the pathogenesis of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). *TARDBP* mutations are postulated to inactivate TDP-43 functions, leading to impaired RNA processing. However, it has not been fully examined how mutant TDP-43 affects global RNA regulation, especially in human cell models. Here, we examined global RNA processing in forebrain cortical neurons derived from human induced pluripotent stem cells (iPSCs) with a pathogenic *TARDBP* mutation encoding the TDP-43^{K263E} protein. In neurons expressing mutant TDP-43, we detected disrupted RNA regulation, including global changes in gene expression, missplicing, and aberrant polyadenylation, all of which were highly similar to those induced by TDP-43 knock-down. This mutation-induced TDP-43 LOF was not because of the cytoplasmic mislocalization of TDP-43. Intriguingly, in nonneuronal cells, including iPSCs and neural progenitor cells (NPCs), we did not observe impairments in RNA processing, thus indicating that the K263E mutation results in neuron-specific LOF of TDP-43. This study characterizes global RNA processing impairments induced by mutant TDP-43 and reveals the unprecedented cell type specificity of TDP-43 LOF in ALS/FTLD pathogenesis.

Key words: amyotrophic lateral sclerosis; frontotemporal lobar degeneration; induced pluripotent stem cell; TDP-43

Significance Statement

Altered RNA metabolism induced by transactivating response element DNA-binding protein of 43 kDa (TDP-43) loss-of-function (LOF) has been suggested to be a core disease mechanism in amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). Pathogenic mutations in *TARDBP* cause ALS and FTLD; thus, these mutations are considered to induce TDP-43 LOF. Here, using a human induced pluripotent stem cell (iPSC)-based model, we found that the pathogenic K263E mutation disrupted RNA processing events in neurons, reflecting TDP-43 LOF. Interestingly, in contrast to neuronal cultures, mutated TDP-43 remained functional in nonneuronal cells, including iPSCs and neural progenitors. This finding indicates an unprecedented neuron-specific LOF of TDP-43 that potentially accounts for the CNS-selective lesions in patients with ALS/FTLD.

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder caused by the degeneration of upper and lower motor neurons, leading to a progressive loss of motor function. Frontotemporal lobar degeneration (FTLD) is the second most common type of dementia after Alzheimer's disease and is characterized by progressive atrophy in the frontal and temporal lobes and by personality and behavioral changes. Transactivating response element DNA-binding protein of 43 kDa (TDP-43) was identified as a major component of pathologic inclusions found in the vast majority of patients with ALS and ~50% of patients with FTLD (Neumann et al., 2006). Moreover, previous studies reported mutations in the *TARDBP* gene encoding TDP-43 in patients with ALS and FTLD (Kabashi et al., 2008; Kovacs et al., 2009). These findings highlight a central role for TDP-43 in ALS/FTLD pathogenesis (S.C. Ling et al., 2013).

TDP-43 has various functions in RNA metabolism, including splicing, polyadenylation, transport, translation, and miRNA synthesis (S.C. Ling et al., 2013). Accumulating evidence suggests that TDP-43 loss-of-function (LOF) underlies the pathomechanism of ALS/FTLD. One particular target of TDP-43 LOF is RNA missplicing. TDP-43 LOF induces the cryptic splicing of a distinct group of genes, such as *STMN2*, which is also observed in patients with ALS/FTLD (Klim et al., 2019; Melamed et al., 2019; Prudencio et al., 2020). Notably, the pathogenesis of TDP-43 LOF is sometimes human-specific, requiring study in human cell models; for example, cryptic splicing of *STMN2* was not conserved in mouse models (J.P. Ling et al., 2015). Therefore, human induced pluripotent stem cell (iPSC)-based disease models have great potential (Okano and Morimoto, 2022). However, TDP-43 knock-down has been mainly used for the characterization of RNA processing impairments in human iPSC-based disease models (Klim et al., 2019; Prudencio et al., 2020), and the link between pathogenic *TARDBP* mutations and RNA regulation has not been fully explored. Another unanswered question is the cell type specificity of the effect of *TARDBP* mutations. As TDP-43 is globally expressed in various organs, including the brain, liver, lung, and kidney (Sephton et al., 2010), it is mysterious why these mutations result in ALS and FTLD, which selectively affect the CNS. Thus, the potential cell type specificity of the effect of the mutant TDP-43 protein on RNA regulation should be investigated.

In this study, using human iPSC-based models, we examined the K263E variant of TDP-43, which was initially identified in a patient with FTLD (Kovacs et al., 2009), with a disrupted RNA binding capacity (H.J. Chen et al., 2019). Similar to TDP-43 knock-down, TDP-43^{K263E} affected various RNA processing machineries in iPSC-derived neurons. These impaired RNA processes included intron splicing and 3′ polyadenylation regulation. In contrast to neurons, however, iPSCs and neural progenitor cells (NPCs) expressing TDP-43^{K263E} did not mimic TDP-43 LOF. This study characterizes global RNA processing impairments induced by mutant TDP-43 and indicates that mutant TDP-43 exhibits LOF in a cell type-specific manner.

Materials and Methods

Culture of undifferentiated iPSCs

Human iPSCs (201B7; Takahashi et al., 2007) were maintained in StemFit AK02 N medium (Ajinomoto). Cells were seeded at a density of 1.5×10^4 cells/well in an iMatrix 511 (Laminin511E8; Wako)-treated six-well plate; 10 μM Y27632 (Nacalai) was only added for the first day. Culture media were changed every other day.

Generation of TDP43^{K263E} isogenic iPSCs

CRISPR-Cas9-mediated homologous recombination was performed by GenAhead Bio. The sgRNA sequences targeting *TARDBP* were as follows: 5′-ATTGTGCTTAGGTTCGGCAT-3′ and 5′-AATAGACAGTTAGAAAGAAG-3′. Both sequences were synthesized and ligated after the U6 promoter. The double-stranded targeting vector (TV) harboring the K263E mutation (c.A > G) together with the silent mutations at S258 (c.C > G) and S273 (c.A > T and c.G > C) flanked by 1-kbp homology arms was assembled using the NEBuilder HiFi DNA Assembly kit (New England BioLabs). iPSCs were electroporated with the Cas9 (D10A) expression vector, both sgRNA vectors, and TV at a ratio of 1:1:2 and verified as containing the biallelic K263E mutant using Sanger sequencing.

Neuronal induction

Neuronal induction of iPSCs was performed using a previously described method (Sato et al., 2021), with slight modifications. Briefly, iPSCs were seeded on iMatrix511-coated 12-well plates at a density of 1.0–1.5×10^5 cells/well in StemFit AK02 N medium. After 3 d, neural induction was initiated by changing the medium to neural induction medium [consisting of Advanced DMEM/F-12 (Thermo Fisher Scientific), 2% B27 supplement (–vitamin A; Thermo Fisher Scientific)] with 150 nM LDN193189 (StemRD), 5 μM SB431542 (Tocris), and 3 μM IWR1e (Calbiochem). On day 6, the cells were dissociated into single cells using Accutase (Nacalai) and

seeded onto poly-L-ornithine-coated and laminin-coated 12-well plates at a 1:1–1:2 ratio. On day 12, the cells were dissociated again, and 8×105 cells seeded in each well of poly-L-ornithine-coated and laminin-coated six-well plates and cultured in neuronal medium [Advanced DMEM/F-12, 2% B27 supplement, 200 μM ascorbic acid (Sigma), and 200 μM dbcAMP (Sigma)] with 20 μM DAPT (Sigma). On day 18, DAPT was removed, and 10 ng/ml BDNF (Alomone Labs), 10 ng/ml GDNF (Alomone Labs), and 1 μM PD0332991 (Sigma) were added. The medium was changed every 3 d.

RNA sequencing

Total RNA was isolated from iPSCs and NPCs on day 12 and from neurons on day 36 with the RNeasy Mini kit (QIAGEN) with DNase I treatment. The quality of RNA (RNA integrity number; RIN) was assessed by Agilent 2100 Bioanalyzer (Agilent). The indexed cDNA libraries were prepared using the TruSeq stranded mRNA Library Preparation kit (Illumina) and sequenced using a NovaSeq6000 (Illumina) to obtain 150-bp paired-end reads at Macrogen. RNA-seq datasets of TDP-43 knock-down experiments from previous reports Klim et al., 2019; Melamed et al., 2019) were downloaded via the NCBI Sequence Read Archive (accession numbers SRR8083864-8, SRR8083871-75, SRR8083878-81, and SRR8144907-12). The RNA-seq dataset of ESCs with TDP-43 knock-down (Modic et al., 2019) was kindly provided by M. Modic (Francis Crick Institute). Raw fastq files were trimmed to remove low-quality bases and adapters using fastp (S. Chen et al., 2018) and were processed for further analyses.

Gene expression profiling

Salmon (Patro et al., 2017) was used to generate the TPM using the transcript index from the reference GRCh38 genome annotation (GENCODE release 33) to quantify gene expression levels. We identified differentially expressed genes (DEGs) using the DESeq2 suite of bioinformatics tools (Love et al., 2014) with a cutoff of 0.05 for Benjamini–Hochberg adjusted p values and a cutoff of 0.25 for the log2 fold change ratio. Principal component analysis (PCA) was performed using vst transformation of estimated counts based on intersections between DEG lists of two independent TDP-43 knock-down experiments from the studies by Klim et al. (2019) and Melamed et al. (2019). Our RNA-seq data were projected onto this PCA.

Read alignment to the genome and alternative splicing analyses

HISAT2 (D. Kim et al., 2019) was used to map sequencing reads to the human GRCh38 genome. Coverage tracks were visualized with Integrative Genomics Viewer (IGV). Counts at individual exons were calculated from the HISAT2-aligned data using featureCounts (Liao et al., 2014) with the gene annotation from Ensembl (release 104), and the differential exon usage analysis was performed using DEXseq (Anders et al., 2012) with a cutoff of 0.01 for Benjamini–Hochberg adjusted p values. Differentially spliced intron

clusters were analyzed from the HISAT2-aligned data without existing isoform annotations by LeafCutter (Li et al., 2018). Briefly, splice junction reads were extracted with RegTools (Cotto et al., 2021) using a minimum of 6 bp as an anchor on each side of the junction. Junctions from each sample were then clustered using leafcutter_cluster_regtools_py3.py (minclureads = 10). Differential intron splicing was calculated using leafcutter_ds.R. In addition to LeafCutter analysis, we also evaluated splice variants of *UNC13A* and *STMN2* by filtering reads that span the junctions between normal exons and cryptic exons as previously described (Ma et al., 2022). Junction spanning reads were filtered and quantified using junction_spanning_reads.sh.

Alternative polyadenylation analysis

QAPA (Ha et al., 2018) was used to quantify TPM for individual alternative polyadenylation sites (PASs) from RNA-seq data. The 3′ untranslated region (UTR) sequence was extracted from GRCh38 genome by qapa fasta based on the precomplied annotation available on the QAPA GitHub page (https://github.com/morrislab/qapa). Salmon index was prepared using this 3′ UTR sequence. 3′ UTR isoform usage was then quantified using "salmon quant" and "qapa quant." PASs with TPMs >5 were retained for further analysis. For each gene, gene-level relative PAS usage was summarized using a metric Ψ (Goering et al., 2021). Each PAS within a gene was assigned a value, m, which is defined as its position within this proximal-to-distal ordering, beginning with 1, to calculate Ψ. Each gene was also assigned a value, n, which is defined as the number of distinct PASs that it contains. The expression level of each PAS (TPM_m) was evaluated, and gene-level PAS usage was summarized using the following formula:

$$\Psi = \frac{\sum_m \left(TPM_m \times \frac{m-1}{n-1} \right)}{\sum_m TPM_m}.$$

Analysis of 3′ end-seq data

The 3′ end-seq datasets from a previous report (Modic et al., 2019) were downloaded from the European Nucleotide Archive (accession numbers ERR1642497 and ERR1642501). Fastp was used for quality and adapter trimming, and sequencing reads were aligned to the human GRCh38 genome using HISAT2. Coverage tracks were visualized using IGV.

Quantitative RT-PCR

cDNA was prepared by using a ReverTraAce qPCR RT kit (Toyobo). The qPCR analysis was performed with TB Green Premix Ex Taq (TAKARA) using a ViiA 7 real-time PCR system (Applied Biosystems) according to the manufacturer's instructions. Values were normalized to *ACTB* levels. Data were analyzed using the comparative (ΔΔCt) method. The primers used for qPCR were as follows:

ACTB, forward 5′-TGAAGTGTGACGTGGACATC-3′, reverse 5′-GGAGGAGCAATGATCTTGAT-3′; *FOXG1*, forward 5′-CCCGTCAATGACTTCGCAGA-3′, reverse 5′-GTCCCGTCGTAAAACTTGGC-3′; *OCT4*, forward 5′-GACAGGGGGAGGGGAGGAGCTAGG-3′, reverse 5′-CTTCCCTCCAACCAGTTGCCCCAAAC-3′; *PAX6,* forward 5′-ACCACACCGGTTTCCTCCTTCACA-3′, reverse 5′-TTGCCATGGTGAAGCTGGGCAT-3′; *TARDBP-ΔIntron7*, forward 5′-TTCATCTCATTTCAAATGTTTATGGAAG-3′, reverse 5′- ATTAACTGCTATGAATTCTTTGCATTCAG-3′; *TUBB3,* forward 5′-ATTTCATCTTTGGTCAGAGTGGGGC-3′, reverse 5′-TGCAGGCAGTCGCAGTTTTCAC-3′; and *UNC13A-CE*, forward 5′-TGGATGGAGAGATGGAACCT-3′, reverse 5′-GGGCTGTCTCATCGTAGTAAAC-3′.

Immunocytochemistry

Neurons cultured until day 36 were fixed with 4% paraformaldehyde for 15 min at room temperature and then washed three times with PBS. After an incubation with blocking buffer (PBS containing 5% normal goat serum and 0.3% Triton X-100) for 30 min at room temperature, the cells were incubated overnight at 4°C with primary antibodies at the following dilutions: TDP-43 (rabbit, Proteintech, 10782-2-AP, 1:200) and TUBB3 (mouse, Sigma, T8660, 1:500). The cells were again washed three times with PBS and incubated with secondary antibodies conjugated to Alexa Fluor 488 or 555 (Life Technologies) and Hoechst 33342 (Dojindo Laboratories) for 1 h at room temperature. After three washes with PBS and one wash with distilled water, the samples were mounted on slides and examined using an LSM-710 confocal laser scanning microscope (Carl Zeiss). Line-scan analysis was performed using ImageJ software. The resulting values were normalized to the maximum intensity. For the analysis of TDP-43 localization, TUBB3 staining was used to determine the cell body as the region of interest, and then Pearson's correlation coefficient was calculated for TDP-43 and Hoechst staining using the Coloc2 plugin.

Data availability

All the sequencing data have been deposited in the NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE195689 (201B7 iPSC) and GSE196144 (otherwise).

Results

Gene expression profile of cortical neurons derived from isogenic iPSCs harboring the TDP-43 K263E mutation

In the present study, we focused on the function of the mutant TDP-43^{K263E} protein (Fig. 1*A*). This mutation of the *TARDBP* gene was identified in a patient with FTLD (Kovacs et al., 2009) and is suggested to reduce the RNA binding capacity of TDP-43 (H.J. Chen et al., 2019). We introduced a homozygous K263E mutation in healthy wild-type human iPSCs using the CRISPR/Cas9 system (Fig. 1*B,C*). Forebrain cortical neurons were generated from control (wild-type) and TDP-43^{K263E} iPSCs by dual SMAD and Wnt inhibition, respectively (Fig. 1*D*; Imaizumi

et al., 2015; Sato et al., 2021). This mutant genotype did not affect the gene expression of markers for pluripotency, neural progenitors, and neurons (Fig. 1*E*). In addition, we found no difference in neuronal induction efficiency (Fig. 1*F*). These data indicate that the pluripotency maintenance and the neuronal induction were equivalent between wild-type and TDP-43^{K263E} iPSCs.

We performed an RNA-seq analysis of these iPSC-derived cortical neurons. Data quality was assured by RIN, Phred quality score, and a uniform mapping rate for all samples (Extended Data Fig. 1-1*A–E*). The global gene expression profile was measured, and the analysis of DEGs identified 550 genes that were significantly differentially expressed between wild-type and TDP-43^{K263E} iPSC-derived neurons (Fig. 1*G*; Extended Data Fig. 1-1*F*). Fold changes of these DEGs were relatively small, but this is consistent with other reports (Klim et al., 2019; Melamed et al., 2019). Among these DEGs, the most significantly altered gene was *STMN2*. As previous studies have reported that *STMN2* is the gene most affected by TDP-43 knock-down (Klim et al., 2019; Melamed et al., 2019), we reanalyzed RNA-seq datasets from these previous studies and compared the effect of TDP-43^{K263E} and TDP-43 knock-down on gene expression. TDP-43 knock-down downregulated *STMN2* expression quite similarly to TDP-43^{K263E} (Fig. 1*H*). PCA grouped TDP-43^{K263E} and TDP-43 knock-down cells into the same cluster along PC2, whereas the data from the study by Melamed and colleagues were grouped separately on PC1 (Fig. 1*I*). When comparing gene expression between TDP-43^{K263E} and TDP-43 knock-down neurons, a strong correlation was observed between TDP-43^{K263E} and TDP-43 knock-down neurons, except for the expression of *TARDBP* (Fig. 1*J,K*). Collectively, TDP-43^{K263E} and TDP-43 knock-down exerted similar effects on gene expression, indicating that K263E corresponds to an LOF mutation.

Characterization of missplicing indicates the similarity between TDP-43^{K263E} and TDP-43 knock-down

TDP-43 plays a pivotal role in RNA processing, and previous reports suggest that *STMN2* loss on TDP-43 knock-down is because of cryptic exon inclusion (Klim et al., 2019; Melamed et al., 2019). Notably, we identified the same cryptic splice events in TDP-43^{K263E} neurons (Fig. 2*A*). This observation suggests that TDP-43^{K263E} induced RNA missplicing; therefore, we analyzed differential exon usage between wild-type and TDP-43^{K263E} neurons using RNA-seq datasets. We identified 299 genes whose exon usages were significantly altered (Fig. 2*B*). We repeated the same exon usage analyses for TDP-43 knock-down datasets, and fold changes in exon usage induced by TDP-43^{K263E} were highly similar to those induced by TDP-43 knock-down (Fig. 2*C*). Notably, both in TDP-43^{K263E} and in TDP-43 knock-down cells, we detected the exclusion of *POLDIP3* exon 3 (Fig. 2*D*), which has previously been associated with TDP-43 deficits (Shiga et al., 2012). Based on these data, RNA missplicing induced by TDP-43^{K263E} results from TDP-43 LOF.

The exon usage analysis is based on existing isoform annotations, but disease-relevant missplicing often occurs in
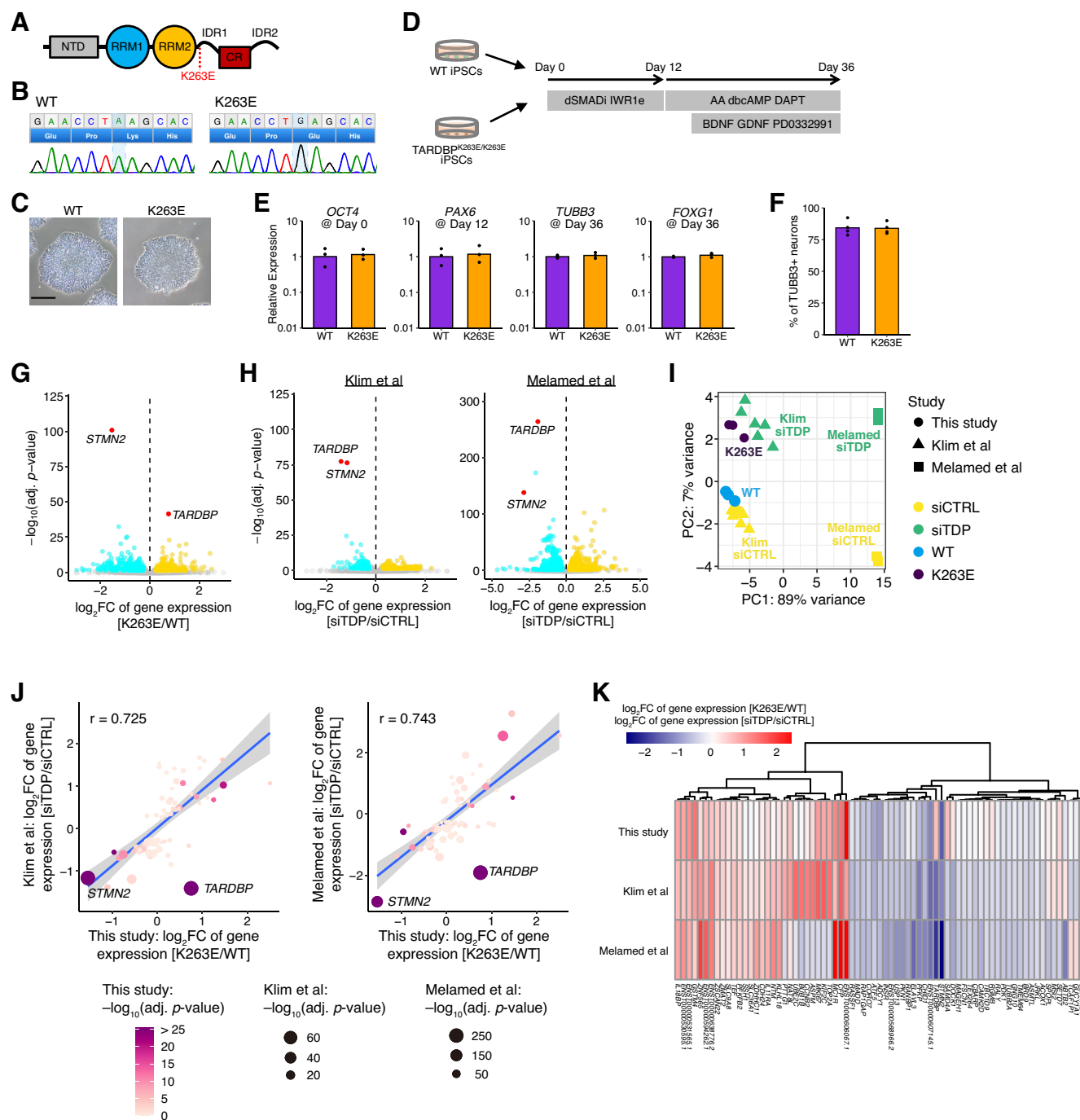
Figure 1. K263E mutation and TDP-43 knock-down exert similar effects on gene expression. **A**, The domain map of TDP-43. The position of the K263E mutation is shown. NTD, N-terminal domain; RPM, RNA recognition motifs; IDR, intrinsically disordered regions; CR, conserved region. **B**, Genomic DNA sequencing of *TARDBP* isogenic iPSC lines. **C**, Bright-field images of wild-type and *TARDBP* isogenic iPSCs. Scale bar: 200 μm. **D**, Schematic of the differentiation of iPSCs into cortical neurons. **E**, Fold changes in the expression of *OCT4* in iPSC, *PAX6* in NPC, and *TUBB3/FOXG1* in neuron cultures (normalized to *ACTB*; n = 3). **F**, Quantification of the number of TUBB3+ neurons at day 36 (n = 4). **G**, Volcano plot showing genes with significantly altered expression in cortical neurons derived from TDP-43^{K263E} iPSCs relative to those from wild-type iPSCs. DEGs were identified using the Wald test with a cutoff of 0.05 for Benjamini–Hochberg adjusted *p* values and a cutoff of 0.25 for the log2 fold change ratio. Quality control of RNA-seq data are shown in Extended Data Figure 1-1. **H**, Volcano plots showing genes with significantly altered expression following TDP-43 knock-down in iPSC-derived neurons and SH-SY5Y cells. The published RNA-seq datasets (Klim et al., 2019; Melamed et al., 2019) were reanalyzed. **I**, PCA of RNA-seq datasets based on DEGs induced by TDP-43 knock-down. **J**, **K**, Scatter plot (**J**) and heatmap (**K**) comparing the fold changes in gene expression induced by TDP-43^{K263E} and TDP-43 knock-down. Analyzed genes were selected as intersections between the DEG lists of two independent TDP-43 knock-down experiments from studies by Klim

continued

and colleagues and Melamed and colleagues without additional filters. Shaded areas indicate 95% confidence intervals. Pearson's correlation coefficients were calculated after excluding the *TARDBP* values.
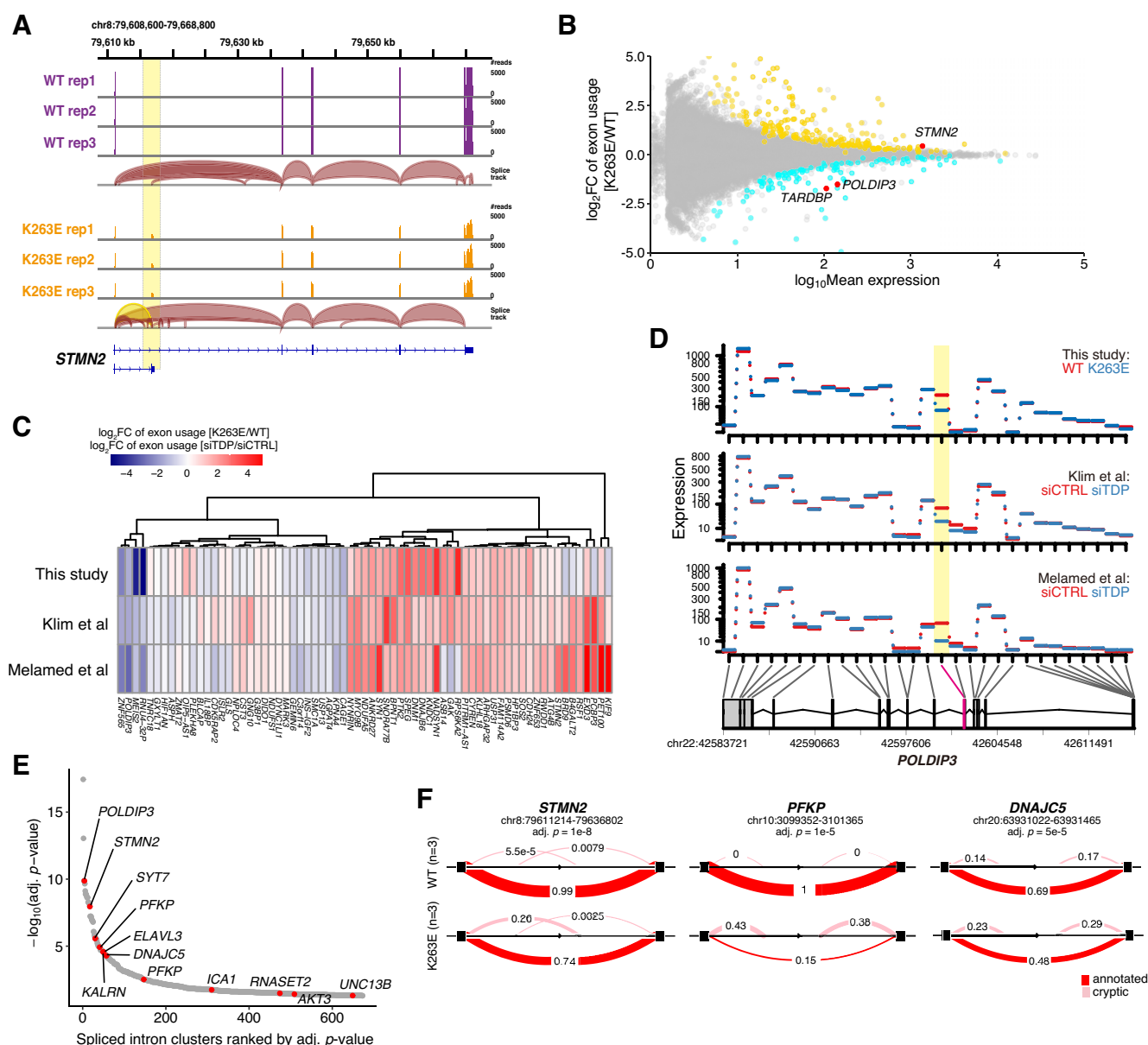
Figure 2. RNA splicing alterations induced by TDP-43[K263E]. **A**, RNA-seq read coverage and splice junctions mapped to the genomic region of *STMN2*. The cryptic exon and splice ribbon from exon 1 to the cryptic exon are highlighted in yellow. **B**, Scatter plot of differential exon usage between TDP-43[K263E] and wild-type iPSC-derived cortical neurons. The expression level at individual exons was quantified. The exon with the most significantly altered expression in each gene is shown. Differentially used exons were identified with a cutoff of 0.01 for Benjamini–Hochberg adjusted *p* values. **C**, Heatmap comparing fold changes in exon usage induced by TDP-43[K263E] and TDP-43 knock-down. Analyzed genes were selected based on intersections between differentially used exon lists of two independent TDP-43 knock-down experiments from studies by Klim and colleagues and Melamed and colleagues. **D**, Visualization of the expression levels of each of the exons of *POLDIP3*. Exons with significantly different usage levels are highlighted in yellow. **E**, Differentially spliced intron clusters identified by LeafCutter analysis with a cutoff of 0.05 for Benjamini–Hochberg adjusted *p* values. Shown are the names of the genes to which the intron clusters belong. **F**, Differentially spliced intron clusters in *STMN2*, *PFKP*, and *DNAJC5* are shown. Exons, black boxes. The thickness of the red band and inserted values represent the proportions of spliced pairs. The evaluation of *UNC13A* splicing is shown in Extended Data Figure 2-1.
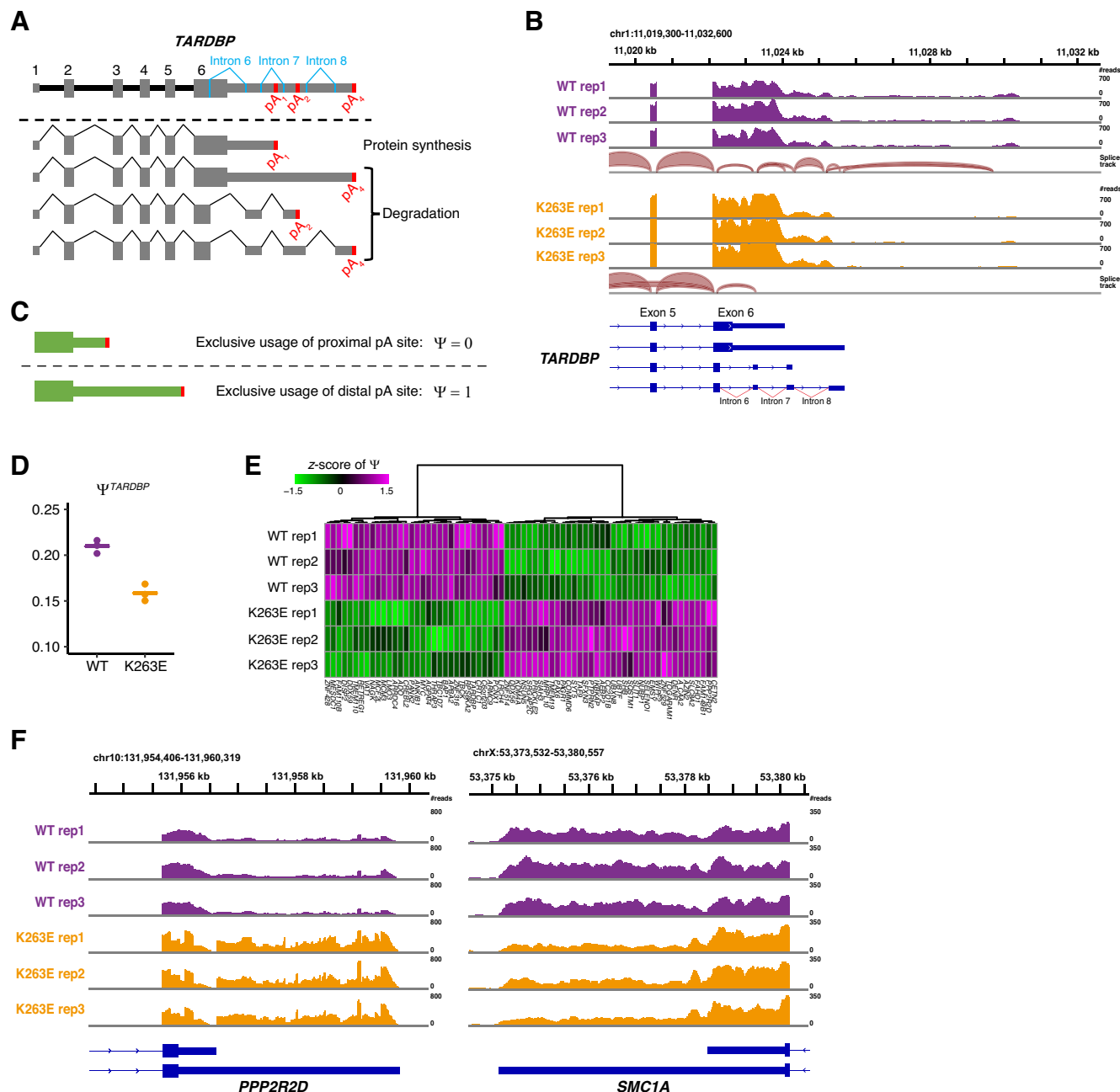
Figure 3. TDP-43[K263E] impairs RNA 3′ end processing. **A**, Schematic diagram of *TARDBP* mRNA isoforms. Alternative polyadenylation and splicing generate several isoforms, most of which are subject to degradation. **B**, RNA-seq read coverage and splice junctions mapped to the genomic region of *TARDBP*. **C**, Ψ as a metric of PAS choice. Genes that exclusively use the proximal PAS are assigned Ψ values of 0, whereas genes that exclusively use distal PAS are assigned Ψ values of 1. **D**, Ψ values for *TARDBP* calculated from RNA-seq datasets in each of three replicates. **E**, Heatmap showing genes with a significant change in Ψ based on a cutoff of 0.01 for *p* values. **F**, RNA-seq read coverage mapped to the genomic regions of *PPP2R2D* and *SMC1A*. 3′ UTRs were differentially transcribed between wild-type and TDP-43[K263E] neurons.

unannotated introns. We performed an annotation-free differential splicing analysis to characterize global changes in intron splicing (Li et al., 2018). We detected cryptic splicing not only in *STMN2* but also in various gene targets, including *PFKP*, *DNAJC5*, *KALRN*, *SYT7*, and *UNC13B*, all of which have been shown to be misspliced in TDP-43 knock-down models (Fig. 2E,F; Polymenidou et al., 2011;

Klim et al., 2019; Brown et al., 2022; Ma et al., 2022). On the other hand, this annotation-free splicing analysis could not detect *UNC13A* cryptic splicing, which has recently been reported as a hallmark of TDP-43 LOF (Brown et al., 2022; Ma et al., 2022). To further evaluate *UNC13A* splicing, sequencing reads spanning the cryptic exon were extracted and quantified. We detected very few spanning
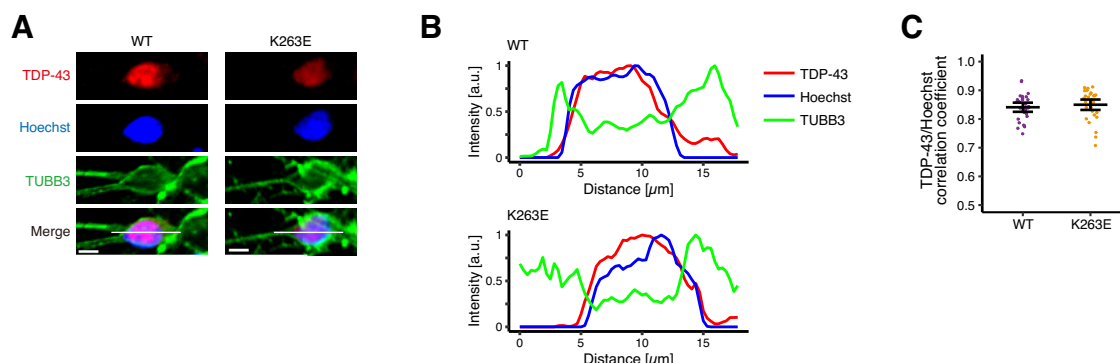
Figure 4. The subcellular localization of TDP-43 is not affected by the K263E mutation. *A*, Representative images of TDP-43 in iPSC-derived cortical neurons. Scale bar: 5 μm. *B*, Line-scan analysis along the solid white lines depicted in *A*. *C*, Pearson's correlation coefficient for TDP-43 immunostaining and Hoechst signals (*n* = 32 cells from 4 independent experiments, mean with 95% confidence interval).

reads in only a subset of samples of TDP-43^{K263E} neurons, whereas there were no spanning reads in wild-type neurons (Extended Data Fig. 2-1*A*–*C*). A similar analysis focusing on *STMN2* cryptic exon robustly detected splicing change by mutant TDP-43 (Extended Data Fig. 2-1*D*–*F*). These results raised two possibilities: one is that TDP-43^{K263E} has little if any effect on *UNC13A* splicing; and the other is that RNA sequencing could not well capture *UNC13A* cryptic exons despite the fact that missplicing occurs. We found that both wild-type and isogenic TDP-43^{K263E} cells harbored homozygous risk SNP/indels (rs12973192, rs12608932, and rs56041637) that increase the efficiency of *UNC13A* cryptic splicing on TDP-43 loss (Extended Data Fig. 2-1*G*; Brown et al., 2022; Ma et al., 2022). Thus, the cryptic splicing of *UNC13A* in these cells is expected to be drastically enhanced by TDP-43 LOF, which supports the second possibility. Indeed, RT-qPCR analysis revealed that *UNC13A* cryptic splicing was reproducibly increased in TDP-43^{K263E} neurons (Extended Data Fig. 2-1*H*). These data suggest that *UNC13A* missplicing was actually occurring, but RNA-seq analysis was not sensitive enough to detect it. The low sensitivity of RNA-seq analysis to detect *UNC13A* missplicing was also supported by the finding that analyses using the data from Klim et al. (2019) and Melamed et al. (2019) similarly detected *UNC13A* missplicing in only a subset of samples (Extended Data Fig. 2-1*I*).

TDP-43 autoregulation through RNA 3′ end processing is impaired by TDP-43^{K263E}

While gene expression profiling and splicing analyses indicated the similarity between TDP-43^{K263E} and TDP-43 knock-down, *TARDBP* expression was upregulated in TDP-43^{K263E} neurons (Fig. 1*D*). *TARDBP* expression was autoregulated by TDP-43 itself; namely, the binding of TDP-43 to its own mRNA changes the processing of the 3′ UTR, finally leading to a decrease in the mRNA level (Fig. 3*A*; Ayala et al., 2011; Eréndira Avendaño-Vázquez et al., 2012; Weskamp and Barmada, 2018). Indeed, we observed decreased splicing events in the 3′ UTR of the

TARDBP mRNA in TDP-43^{K263E} neurons (Fig. 3*B*), suggesting that K263E mutations affect the autoregulatory properties of TDP-43. We focused on alternative polyadenylation, which results in the formation of multiple transcript isoforms with distinct 3′ UTRs, to further investigate 3′ UTR regulation. We quantified the usage level of each PAS from RNA-seq data (Ha et al., 2018), and gene-level PAS usage was summarized as the metric Ψ (Goering et al., 2021). Genes with exclusive usage of the most proximal PAS were assigned Ψ values of 0, whereas genes with exclusive usage of the most distal PAS were assigned Ψ values of 1 (Fig. 3*C*). The use of multiple PASs will result in Ψ values between 0 and 1, depending on the relative usage of individual sites. With this metric Ψ, we found that the 3′ UTR was shortened in TDP-43^{K263E} neurons compared with wild-type neurons (Fig. 3*D*). These results imply that the K263E mutation impairs the 3′ UTR regulation of the *TARDBP* mRNA, leading to the collapse of autoregulation.

Global disruption of 3′ end processing

TDP-43 regulates not only its own mRNA but also a broad range of 3′ end processing events (Rot et al., 2017). Thus, we performed the differential PAS usage analysis with the metric Ψ and identified 70 genes whose PAS usage was significantly altered between wild-type and TDP-43^{K263E} neurons (Fig. 3*E*). The proximal PAS of *PPP2R2D* was preferentially used in wild-type neurons; on the other hand, TDP-43^{K263E} neurons used the lengthened 3′ UTR (Fig. 3*F*). The 3′ UTR of *SMC1A* was shifted similarly, but in the opposite direction (Fig. 3*F*). These shifts are consistent with a recent study that identified genes with PAS switches induced by TDP-43 depletion (Hallegger et al., 2021). Collectively, these results suggest that TDP-43^{K263E} affects RNA 3′ end processing of various genes because of the impaired TDP-43 function.

The subcellular localization of TDP-43 is unaffected

Based on our results, the LOF of TDP-43 has been suggested in TDP-43^{K263E} neurons. TDP-43 functions mainly in the nucleus, and its redistribution from the nucleus to
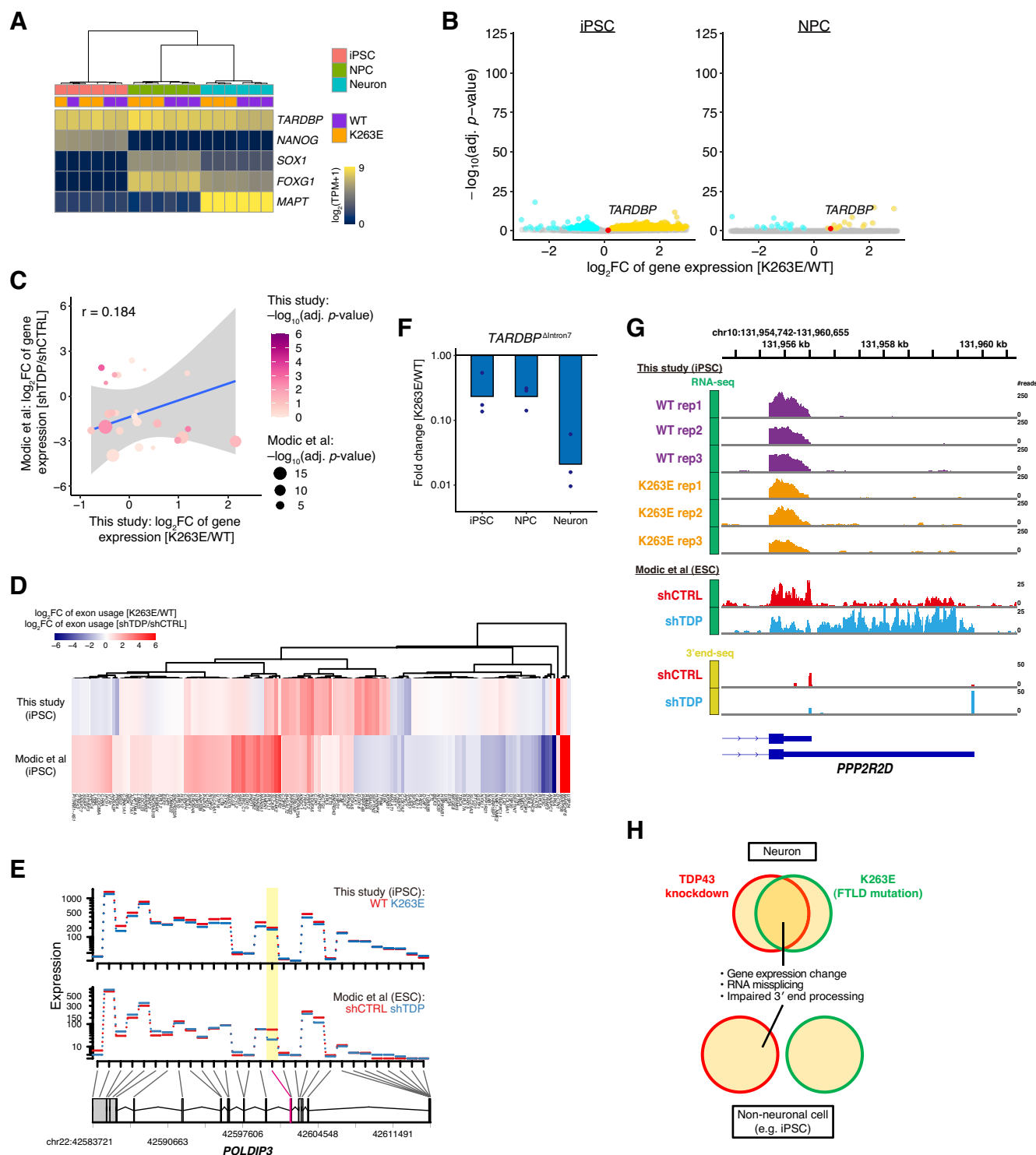
Figure 5. RNA processing impairments induced by TDP-43^{K263E} are not observed in nonneuronal cells. *A*, Heatmap of gene expression in iPSCs, NPCs, and neurons. *B*, Volcano plots showing genes with significantly altered expression between wild-type and TDP-43^{K263E} iPSCs and NPCs. *C*, Scatter plot comparing fold changes in gene expression induced by TDP-43^{K263E} in iPSCs with those induced by TDP-43 knock-down in ESCs (Modic et al., 2019). DEGs in the study by Modic and colleagues (using a cutoff of 0.05 for Benjamini–Hochberg adjusted *p* values and a cutoff of 0.25 for the log2 fold change ratio) were analyzed. Shaded areas indicate 95% confidence intervals. Pearson's correlation coefficients are shown. *D*, Heatmap comparing fold changes in the exon usage induced by TDP-43^{K263E} in iPSCs with those induced by TDP-43 knock-down in ESCs. Differentially used exons in the study by Modic and colleagues (using a cutoff of 0.01 for Benjamini–Hochberg adjusted *p* values) were analyzed. *E*, Visualization of the expression levels of each of the exons of *POLDIP3* in iPSC/ESC cultures. Exons with significantly different usage levels are highlighted in yellow. *F*, Fold changes in the expression of the *TARDBP* isoform with intron 7 splicing in iPSC, NPC, and neuron cultures

---

(normalized to *ACTB*; $n = 3$). **G**, Top panel, RNA-seq read coverage mapped to the genomic region of *PPP2R2D* in wild-type and TDP-43^{K263E} iPSC cultures. Middle panel, RNA-seq read coverage in ESC cultures with TDP-43 knock-down. Bottom panel, 3′ end-seq coverage tracks in ESC cultures with TDP-43 knock-down. **H**, Schematic illustrating neuron-specific effects of the K263E mutation on the function of TDP-43. In neuronal cell cultures, TDP-43^{K263E} phenocopies TDP-43 knock-down, including changes in gene expression, misregulation of RNA splicing, and aberrant 3′ end processing. On the other hand, in nonneuronal cell types, TDP-43^{K263E} does not exhibit LOF phenotypes.

---

the cytoplasm has been recognized as a pathologic hallmark of ALS and FTLD, suggesting that the pathogenic mechanism is associated with the loss of nuclear TDP-43 function (Mackenzie et al., 2010; G. Kim et al., 2020). Thus, we examined whether K263E exerted its LOF effect through subcellular mislocalization. In both wild-type and TDP-43^{K263E} neurons, we observed signals for TDP-43 immunostaining primarily in the nucleus (Fig. 4A,B). Pearson's correlation coefficient revealed a strong correlation between TDP-43 immunostaining and the DNA counterstain in both groups (Fig. 4C). In addition, while intranuclear droplets of TDP-43 are suggested to be precursors of nuclear or cytoplasmic TDP-43 aggregates (Yu et al., 2021), TDP-43 was diffusely distributed within the nucleus in both wild-type and TDP-43^{K263E} neurons (Fig. 4A,B). Therefore, the TDP-43 subcellular or intranuclear localization was not altered by the K263E mutation.

Cell type specificity of disrupted RNA processing induced by TDP-43^{K263E}

We performed RNA-seq of iPSCs and NPCs expressing wild-type or K263E variant TDP-43 to test whether these RNA processing disruptions induced by TDP-43^{K263E} were also present in cell types other than neurons (Extended Data Fig. 1-1A–E). *TARDBP* was expressed at similarly high levels among iPSCs, NPCs, and neurons (Fig. 5A). In contrast to neurons, we did not find a drastic change of *TARDBP* expression in TDP-43^{K263E} iPSCs and NPCs compared with wild-type cells (Fig. 5B). We reanalyzed the RNA-seq data from ESCs with TDP-43 knock-down (Modic et al., 2019), and the comparison of the gene expression profiles showed no correlation between TDP-43^{K263E} iPSCs and TDP-43 knock-down ESCs (Fig. 5C), suggesting that the K263E mutation does not phenocopy TDP-43 knock-down in iPSC cultures. This dissimilarity between TDP-43^{K263E} iPSCs and TDP-43 knock-down ESCs was also evident in the differential exon usage analysis, in which exon usage fold changes induced by TDP-43^{K263E} were quite distinct from those induced by TDP-43 knock-down (Fig. 5D). Remarkably, the exclusion of *POLDIP3* exon 3 was only observed in TDP-43 knock-down ESCs but not in TDP-43^{K263E} iPSCs (Fig. 5E). Moreover, the K263E mutation drastically decreased intron 7 splicing in the 3′ UTR of *TARDBP* only in neurons; in contrast, this splicing event was rarely affected in TDP-43^{K263E} iPSCs and NPCs (Fig. 5F). In the analysis of the 3′ end processing of the *PPP2R2D* mRNA, RNA-seq and 3′ end-seq revealed that the 3′ UTR was lengthened in TDP-43 knock-down ESCs, but TDP-43^{K263E} iPSCs did not exhibit this

PAS switch (Fig. 5G). Taken together, the data from nonneuronal cells, especially iPSCs, indicate that TDP-43^{K263E} did not mimic the RNA processing impairment induced by TDP-43 depletion. This result is the opposite of the close relationship between the K263E mutation and LOF in neurons (Fig. 5H).

Discussion

Human iPSC-based disease models have remarkable potential to clarify disease pathogenesis and to discover effective therapies because of the discrepancy between patients and animal models (Okano and Yamanaka, 2014; Imaizumi and Okano, 2021). Previous studies of human iPSC-based models revealed RNA misprocessing induced by TDP-43 LOF; however, these studies largely used a knock-down strategy for recapitulating TDP-43 LOF (Klim et al., 2019; Prudencio et al., 2020), and the effect of pathogenic mutations of TDP-43 on global RNA regulation has not been extensively investigated. In the present study, we characterized global RNA processing impairments induced by the K263E mutation in iPSC-derived neurons. Consistent with the disrupted RNA binding capacity of TDP-43^{K263E} (H.J. Chen et al., 2019), substantial similarity exists between our mutant model and the knock-down system, indicating that pathogenic TDP-43 mutation induces TDP-43 LOF in neurons. Notably, our mutant model exhibited a defect in TDP-43 autoregulation, which cannot be investigated in a knock-down system.

Previous reports have suggested several distinct but overlapping effects of mutations on TDP-43 functions, including cytoplasmic mislocalization, an increased tendency to aggregate, and a decreased RNA binding capacity (Prasad et al., 2019). Cellular models overexpressing the TDP-43^{K263E} mutant indicate that K263E disrupts the RNA binding capacity of TDP-43 and enhances intranuclear TDP-43 aggregation or liquid shell formation (H.J. Chen et al., 2019; Yu et al., 2021). Although our results are consistent with the reduced RNA binding capacity of TDP-43^{K263E}, we did not observe aggregation or droplets of TDP-43 in our iPSC-derived neurons. This inconsistency is probably because of the difference in TDP-43 expression levels. Excess TDP-43 expression might enhance TDP-43 aggregation, but endogenous TDP-43 expression does not. This finding is supported by a neuropathological study of a patient with FTLD carrying the K263E mutation, in which intranuclear TDP-43 inclusions were not observed in the cerebral cortex (Kovacs et al., 2009).

Additionally, cytoplasmic TDP-43 inclusions were indeed observed in this K263E carrier patient (Kovacs et al., 2009). In contrast to this pathologic observation,

iPSC-derived neurons did not exhibit cytoplasmic mislocalization of TDP-43, consistent with previous reports using iPSC-derived neurons expressing other pathogenic TDP-43 mutations (Klim et al., 2019). On the other hand, cytoplasmic TDP-43 accumulation was observed in neurons directly converted from fibroblasts expressing another pathogenic TDP-43 mutation, N352S (Melamed et al., 2019). A recent study of Huntington's disease suggested that age-related signatures were different between iPSC-derived neurons and directly converted neurons; therefore, pathogenic huntingtin aggregates were detected only in converted neurons but not in iPSC-derived neurons (Victor et al., 2018). This study revealed that the induction of pluripotency erased age marks, while direct neuronal conversion maintained age-related signatures. Therefore, direct neuronal conversion would be beneficial for investigating age-associated phenotypes associated with TDP-43 mutations reflecting more advanced pathology.

Our results strongly suggest the cell type-specific function of TDP-43^{K263E}. TDP-43 LOF was only observed in neuronal cultures, whereas iPSCs and NPCs did not suffer from K263E mutation. The effect of TDP-43 mutation is often tested in nonneuronal cell lines, and comparisons among cell lines have been seldom studied. Thus, the findings for the TDP-43 mutant may need to be revisited in terms of the cell types tested. This cell type specificity is also important for the pathogenesis of ALS/FTLD. CNS-specific lesions in patients with ALS/FTLD do not match the observation that TDP-43 is expressed in various organs in the body (Sephton et al., 2010). Neuron-specific TDP-43 LOF may account for CNS-specific pathology. Although we do not yet know the mechanism of cell type-specific LOF, a neuron-specific cofactor of TDP-43 might have a role. Regardless, this study provides insights into the unprecedented cell type specificity of TDP-43 LOF and provides new opportunities for advancing ALS/FTLD research.

References

Anders S, Reyes A, Huber W (2012) Detecting differential usage of exons from RNA-seq data. Genome Res 22:2008–2017.

Ayala YM, De Conti L, Avendaño-Vázquez SE, Dhir A, Romano M, D'Ambrogio A, Tollervey J, Ule J, Baralle M, Buratti E, Baralle FE (2011) TDP-43 regulates its mRNA levels through a negative feedback loop. EMBO J 30:277–288.

Brown AL, et al. (2022) TDP-43 loss and ALS-risk SNPs drive missplicing and depletion of UNC13A. Nature 603:131–137.

Chen HJ, Topp SD, Hui HS, Zacco E, Katarya M, McLoughlin C, King A, Smith BN, Troakes C, Pastore A, Shaw CE (2019) RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. Brain 142:3753–3770.

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.

Cotto KC, Feng Y-Y, Ramu A, Skidmore ZL, Kunisaki J, Richters M, Freshour S, Lin Y, Chapman WC, Uppaluri R, Govindan R, Griffith OL, Griffith M (2021) RegTools: integrated analysis of genomic and transcriptomic data for the discovery of splicing variants in cancer. bioRxiv 436634.

Eréndira Avendaño-Vázquez S, Dhir A, Bembich S, Buratti E, Proudfoot N, Baralle FE (2012) Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. Genes Dev 26:1679–1684.

Goering R, Engel KL, Gillen AE, Fong N, Bentley DL, Taliaferro JM (2021) LABRAT reveals association of alternative polyadenylation with transcript localization, RNA binding protein expression, transcription speed, and cancer survival. BMC Genomics 22:476.

Ha KCH, Blencowe BJ, Morris Q (2018) QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. Genome Biol 19:45.

Hallegger M, Chakrabarti AM, Lee FCY, Lee BL, Amalietti AG, Odeh HM, Copley KE, Rubien JD, Portz B, Kuret K, Huppertz I, Rau F, Patani R, Fawzi NL, Shorter J, Luscombe NM, Ule J (2021) TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell 184:4680–4696.e22.

Imaizumi K, Okano H (2021) Modeling neurodevelopment in a dish with pluripotent stem cells. Dev Growth Differ 63:18–25.

Imaizumi K, Sone T, Ibata K, Fujimori K, Yuzaki M, Akamatsu W, Okano H (2015) Controlling the regional identity of hPSC-derived neurons to uncover neuronal subtype specificity of neurological disease phenotypes. Stem Cell Reports 5:1010–1022.

Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, Bouchard JP, Lacomblez L, Pochigaeva K, Salachas F, Pradat PF, Camu W, Meininger V, Dupre N, Rouleau GA (2008) TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet 40:572–574.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915.

Kim G, Gautier O, Tassoni-Tsuchida E, Ma XR, Gitler AD (2020) ALS genetics: gains, losses, and implications for future therapies. Neuron 108:822–842.

Klim JR, Williams LA, Limone F, Guerra San Juan I, Davis-Dusenbery BN, Mordes DA, Burberry A, Steinbaugh MJ, Gamage KK, Kirchner R, Moccia R, Cassel SH, Chen K, Wainger BJ, Woolf CJ, Eggan K (2019) ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. Nat Neurosci 22:167–179.

Kovacs GG, Murrell JR, Horvath S, Haraszti L, Majtenyi K, Molnar MJ, Budka H, Ghetti B, Spina S (2009) TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. Mov Disord 24:1842–1847.

Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, Pritchard JK (2018) Annotation-free quantification of RNA splicing using LeafCutter. Nat Genet 50:151–158.

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930.

Ling JP, Pletnikova O, Troncoso JC, Wong PC (2015) TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science 349:650–655.

Ling SC, Polymenidou M, Cleveland DW (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron 79:416–438.

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550–521.

Ma XR, Zheng XL, Zheng CY, Hu YT, Qin H, Chen JH, Xu QF, Liang CF (2022) TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A. Nature 603:124–130.

Mackenzie IR, Rademakers R, Neumann M (2010) TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. Lancet Neurol 9:995–1007.

Melamed Z, López-Erauskin J, Baughn MW, Zhang O, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, Ohkubo T, Rodriguez M, Lin N, Wu D, Bennett CF, Rigo F, Da Cruz S, Ravits J, Lagier-Tourenne C, Cleveland DW (2019) Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. Nat Neurosci 22:180–190.

Modic M, Grosch M, Rot G, Schirge S, Lepko T, Yamazaki T, Lee FCY, Rusha E, Shaposhnikov D, Palo M, Merl-Pham J,

Cacchiarelli D, Rogelj B, Hauck SM, von Mering C, Meissner A, Lickert H, Hirose T, Ule J, Drukker M (2019) Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. Mol Cell 74:951–965.e13.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314:130–133.

Okano H, Yamanaka S (2014) iPS cell technologies: significance and applications to CNS regeneration and disease. Mol Brain 7:22.

Okano H, Morimoto S (2022) iPSC-based disease modeling and drug discovery in cardinal neurodegenerative disorders. Cell Stem Cell 29:189–208.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419.

Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling SC, Sun E, Wancewicz E, Mazur C, Kordasiewicz H, Sedaghat Y, Donohue JP, Shiue L, Bennett CF, Yeo GW, Cleveland DW (2011) Long pre-mRNA depletion and RNA mis-splicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci 14:459–468.

Prasad A, Bharathi V, Sivalingam V, Girdhar A, Patel BK (2019) Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis. Front Mol Neurosci 12:25.

Prudencio M, et al. (2020) Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia. J Clin Invest 130:6080–6092.

Rot G, Wang Z, Huppertz I, Modic M, Lenče T, Hallegger M, Haberman N, Curk T, von Mering C, Ule J (2017) High-resolution RNA maps suggest common principles of splicing and polyadenylation regulation by TDP-43. Cell Rep 19:1056–1067.

Sato T, Imaizumi K, Watanabe H, Ishikawa M, Okano H (2021) Generation of region-specific and high-purity neurons from human feeder-free iPSCs. Neurosci Lett 746:135676.

Sephton CF, Good SK, Atkin S, Dewey CM, Mayer P, Herz J, Yu G (2010) TDP-43 is a developmentally regulated protein essential for early embryonic development. J Biol Chem 285:6826–6834.

Shiga A, Ishihara T, Miyashita A, Kuwabara M, Kato T, Watanabe N, Yamahira A, Kondo C, Yokoseki A, Takahashi M, Kuwano R, Kakita A, Nishizawa M, Takahashi H, Onodera O (2012) Alteration of POLDIP3 splicing associated with loss of function of TDP-43 in tissues affected with ALS. PLoS One 7:e43120.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872.

Victor MB, Richner M, Olsen HE, Lee SW, Monteys AM, Ma C, Huh CJ, Zhang B, Davidson BL, Yang XW, Yoo AS (2018) Striatal neurons directly converted from Huntington's disease patient fibroblasts recapitulate age-associated disease phenotypes. Nat Neurosci 21:341–352.

Weskamp K, Barmada SJ (2018) TDP43 and RNA instability in amyotrophic lateral sclerosis. Brain Res 1693:67–74.

Yu H, Lu S, Gasior K, Singh D, Vazquez-Sanchez S, Tapia O, Toprani D, Beccari MS, Yates JR, Da Cruz S, Newby JM, Lafarga M, Gladfelter AS, Villa E, Cleveland DW (2021) HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. Science 371:eabb4309.