1. TITLE
Pathogenic mutation of TDP-43 impairs RNA processing in a cell type-specific manner: implications for the pathogenesis of ALS/FTLD

2. ABBREVIATED TITLE
Neuron-specific RNA misprocessing by TDP-43 mutation

3. AUTHORS
Kent Imaizumi¹, Hirosato Ideno¹, Tsukika Sato¹, Satoru Morimoto¹, and Hideyuki Okano¹.
¹Department of Physiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku, Tokyo 160-8582, Japan

4. AUTHOR CONTRIBUTIONS
KI, SM, and HO designed research; KI, HI, and TS performed research; KI analyzed data; KI and HO wrote the paper.

5. CORRESPONDENCE
Correspondence should be addressed to Dr. Hideyuki Okano, Department of Physiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku, Tokyo 160-8582, Japan. E-mail: hidokano@a2.keio.jp.

6. Number of figures: 5
7. Number of tables: 0
8. Number of multimedia: 0
9. Number of words for Abstract: 204
10. Number of words for Significance Statement: 110
11. Number of words for Introduction: 474
12. Number of words for Discussion: 605

13. ACKNOWLEDGMENTS
We are grateful to M. Yano (Niigata University) for his feedback on the manuscript and discussions; T. Matsuo (Takeda Pharmaceutical Company) and T. Akiyama (Stanford University) for discussions; S. Yamanaka (Kyoto University) for 201B7 iPSCs; M. Modic (Francis Crick Institute) for kindly providing RNA-seq dataset of ESCs with TDP-43 knockdown; GenAhead Bio for generating isogenic TDP-43K263E iPSCs; and all members of the H.O. laboratory for encouragement and kind support.
14. CONFLICT OF INTEREST
H.O. serves as a paid scientific advisor to SanBio Co. Ltd. and K Pharma Inc. The remaining authors declare no competing financial interests.

15. FUNDING SOURCES
This work was supported by funding from Takeda Pharmaceutical Company, Ltd, Japan, Japan Agency for Medical Research and Development (AMED) (Grant Number 19bm0804003, 20bm0804003, and 21bm0804003), and JSPS Grant-in-Aid for Scientific Research (Grant Number 20H00485) to H.O., and Grant-in-Aid for Scientific Research on Innovative Areas “Singularity Biology (No. 8007)” (Grant Number 21H00438) to K.I.
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-43K263E 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 knockdown. This mutation-induced TDP-43 LOF was not due to the cytoplasmic mislocalization of TDP-43. Intriguingly, in nonneuronal cells, including iPSCs and neural progenitor cells, 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.

SIGNIFICANCE STATEMENT

Altered RNA metabolism induced by 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 central nervous system-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 pathological inclusions found in the vast majority of patients with ALS and approximately 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 (Ling et al., 2013).

TDP-43 has various functions in RNA metabolism, including splicing, polyadenylation, transport, translation, and microRNA synthesis (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 (Ling et al., 2015). Therefore, human induced pluripotent stem cell (iPSC)-based disease models have great potential (Okano and Morimoto, 2022). However, TDP-43 knockdown 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 central nervous system (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 (Chen et al., 2019). Similar to TDP-43 knockdown, 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 $1.5 \times 10^4$ cells/well in an iMatrix 511 (Laminin511E8; Wako)-treated 6-well plate. 10 μM Y27632 (Nacalai) was only added for the first day. Culture media were changed every other day.

Generation of TDP43K263E 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 $1.0–1.5 \times 10^5$ cells/well in StemFit AK02 N medium. After three days, neural induction was initiated by changing the medium to neural induction medium (consisting of Advanced DMEM/F-12 (Thermo Fisher), 2% B27 supplement (–vitamin A) (Thermo Fisher)) 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- and laminin-coated 12-well plates at a 1:1–1:2 ratio. On Day 12, the cells were dissociated again, and $8 \times 10^3$ cells seeded in each well of poly-L-ornithine- and laminin-coated 6-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 days.

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 knockdown 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 knockdown (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 (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 knockdown 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 (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’ (mininreads = 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. 3’ UTR sequence was extracted from GRCh38 genome by ‘qapa fasta’ based on the pre-complied annotation available on the QAPA GitHub.
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 greater than 5 were retained for further analysis. For each gene, gene-level relative PAS usage was summarized using a metric \( \Psi \) (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 \( \Psi \).

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:

\[
\Psi = \frac{\sum_m (TPM_m \times \frac{m-1}{n-1})}{\sum_m TPM_m}
\]

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 (\( \Delta \Delta Ct \)) method. The primers used for qPCR were as follows: ACTB, forward 5′-TGAAGTGTGACGTGGACATC-3′, reverse 5′-GGAGGAGCAATGATCTTTG-3′; FOXG1, forward 5′-CCCCTGAATGACTTCGA-3′, reverse 5′-GTCCCGTCGTGTTTGGCTGGGGGAGAGCTAGG-3′; OCT4, forward 5′-GACAGGGGGAGGGGAGGATCAGGGGAGAGCTAGG-3′, reverse 5′-CTGCCCTCACCAGTTCGCCAAAAC-3′; PAX6, forward 5′-ACCACACCAGGTTTCTCCTGACA-3′, reverse 5′-TTGAGATGCTGATGAGC-3′; TARDBP-DIntron7, forward 5′-TTCCATTTTCAATTTGGAC-3′, reverse 5′-ATTAACTGCTATTTCTTTG-3′; TUBB3, forward 5′-ATTTATTTTACGAGTAGG-3′, reverse 5′-TCGAGAGTGGAGATGACCT-3′; and UNC13A-CE, forward 5′-TGAGATGAGATGGAACCT-3′, reverse 5′-GGGTGTCTCATCGATGAAAC-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<sup>K263E</sup> protein (Fig. 1A). 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 (Chen et al., 2019). We introduced a homozygous K263E mutation in healthy wild-type human iPSCs using the CRISPR/Cas9 system (Fig. 1B, C). Forebrain cortical neurons were generated from control (wild-type) and TDP-43<sup>K263E</sup> iPSCs by dual SMAD and Wnt inhibition, respectively (Fig. 1D) (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. 1E). In addition, we found no difference in neuronal induction efficiency (Fig. 1F). These data indicate that the pluripotency maintenance and the neuronal induction were equivalent between wild-type and TDP-43<sup>K263E</sup> iPSCs.

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

Characterization of missplicing indicates the similarity between TDP-43<sup>K263E</sup> and TDP-43 knockdown

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

The exon usage analysis is based on existing isoform annotations, but disease-relevant missplicing often occurs in 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 knockdown 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 reads in only a subset of samples of TDP-43\textsuperscript{K263E} neurons, whereas there were no spanning reads in wild-type neurons (Fig. 2–2A–C). A similar analysis focusing on STMN2 cryptic exon robustly detected splicing change by mutant TDP-43 (Fig. 2–2D–F). These results raised two possibilities: one is that TDP-43\textsuperscript{K263E} has little if any effect on UNC13A splicing; and the other is that RNA sequencing could not well capture UNC13A cryptic exons in spite of the fact that missplicing occurs. We found that both wild-type and isogenic TDP-43\textsuperscript{K263E} cells harbored homozygous risk SNP/indels (rs12973192, rs12608932, and rs56041637) that increase the efficiency of UNC13A cryptic splicing upon TDP-43 loss (Fig. 2–2G) (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\textsuperscript{K263E} neurons (Fig. 2–2H). 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 (Fig. 2–2I).

TDP-43 autoregulation through RNA 3′ end processing is impaired by TDP-43\textsuperscript{K263E}
While gene expression profiling and splicing analyses indicated the similarity between TDP-43<sup>K263E</sup> and TDP-43 knockdown, TARDBP expression was upregulated in TDP-43<sup>K263E</sup> neurons (Fig. 1D). TARDBP expression was autoregulated by TDP-43 itself; namely, the binding of TDP-43 to its own mRNA changes the processing of the 3′ untranslated region (UTR), finally leading to a decrease in the mRNA level (Fig. 3A) (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<sup>K263E</sup> neurons (Fig. 3B), 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 polyadenylation site (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. 3C). 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<sup>K263E</sup> neurons compared with wild-type neurons (Fig. 3D). 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<sup>K263E</sup> neurons (Fig. 3E). The proximal PAS of PPP2R2D was preferentially used in wild-type neurons; on the other hand, TDP-43<sup>K263E</sup> neurons used the lengthened 3′ UTR (Fig. 3F). The 3′ UTR of SMC1A was shifted similarly, but in the opposite direction (Fig. 3F). 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<sup>K263E</sup> affects RNA 3′ end processing of various genes due to 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<sup>K263E</sup> neurons. TDP-43 functions mainly in the nucleus, and its redistribution from the nucleus to the cytoplasm has been recognized as a pathological hallmark of ALS and FTLD, suggesting that the pathogenic mechanism is associated with the loss of nuclear TDP-43 function (Mackenzie et al., 2010; Kim et al., 2020). Thus, we examined whether K263E exerted its LOF effect through subcellular mislocalization. In both wild-type and TDP-43<sup>K263E</sup> neurons, we observed signals for TDP-43 immunostaining primarily in the nucleus (Fig.
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-43K263E 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-43K263E**

We performed RNA-seq of iPSCs and neural progenitor cells (NPCs) expressing wild-type or K263E variant TDP-43 to test whether these RNA processing disruptions induced by TDP-43K263E were also present in cell types other than neurons (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-43K263E iPSCs and NPCs compared with wild-type cells (Fig. 5B). We reanalyzed the RNA-seq data from ESCs with TDP-43 knockdown (Modic et al., 2019), and the comparison of the gene expression profiles showed no correlation between TDP-43K263E iPSCs and TDP-43 knockdown ESCs (Fig. 5C), suggesting that the K263E mutation does not phenocopy TDP-43 knockdown in iPSC cultures. This dissimilarity between TDP-43K263E iPSCs and TDP-43 knockdown ESCs was also evident in the differential exon usage analysis, in which exon usage fold changes induced by TDP-43K263E were quite distinct from those induced by TDP-43 knockdown (Fig. 5D).

Remarkably, the exclusion of POLDIP3 exon 3 was only observed in TDP-43 knockdown ESCs but not in TDP-43K263E 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-43K263E 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 knockdown ESCs, but TDP-43K263E iPSCs did not exhibit this PAS switch (Fig. 5G). Taken together, the data from nonneuronal cells, especially iPSCs, indicate that TDP-43K263E 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
knockdown 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-43K263E (Chen et al., 2019), substantial similarity exists between our mutant model and the
knockdown 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 knockdown 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-43K263E mutant
indicate that K263E disrupts the RNA binding capacity of TDP-43 and enhances intranuclear TDP-43
aggregation or liquid shell formation (Chen et al., 2019; Yu et al., 2021). Although our results are
consistent with the reduced RNA binding capacity of TDP-43K263E, we did not observe aggregation or
droplets of TDP-43 in our iPSC-derived neurons. This inconsistency is probably due to the difference in
TDP-43 expression levels. Excess TDP-43 expression might enhance TDP-43 aggregation, but
endoogenous 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 pathological 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\textsuperscript{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


14

a (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


Kim G, Gautier O, Tassoni-Tsuchida E, Ma XR, Gitler AD (2020) ALS Genetics: Gains, Losses, and


Kim 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


Kovacs GG, Murrell JR, Horvath S, Haraszi L, Majtenyi K, Molnar MJ, Budka H, Ghetti B, Spina S

(2009) TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and


Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, Pritchard JK (2018) Annotation-free


is compromised in ALS-FTD. Science 349:650–655.

Ling S-C, Polymenidou M, Cleveland DW (2013) Converging Mechanisms in ALS and FTD: Disrupted


Ma XR et al. (2022) TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A. Nature

603:124–130.


Melamed Z et al. (2019) Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of


Modic M et al. (2019) Cross-Regulation between TDP-43 and Paraspeckles Promotes


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,

degeneration and amyotrophic lateral sclerosis. Science 314:130–133.


Figure 1. K263E mutation and TDP-43 knockdown 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-43K263E iPSCs relative to those from wild-type iPSCs. Differentially expressed genes 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 is shown in Extended Data Fig. 1–1.

H, Volcano plots showing genes with significantly altered expression following TDP-43 knockdown 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 differentially expressed genes induced by TDP-43 knockdown.

J, K, Scatter plot (J) and heatmap (K) comparing the fold changes in gene expression induced by TDP-43K263E and TDP-43 knockdown. Analyzed genes were selected as intersections between the DEG lists of two independent TDP-43 knockdown experiments from studies by Klim et al. and Melamed et al. 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-43K263E

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-43K263E 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-43K263E and TDP-43 knockdown. Analyzed genes were selected based on intersections between differentially used exon lists of two independent TDP-43 knockdown experiments from studies by Klim et al. and Melamed et al.
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 Fig. 2–1.

Figure 3. TDP-43K263E 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 3 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-43K263E neurons.

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).

Figure 5. RNA processing impairments induced by TDP-43K263E 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-43K263E iPSCs and NPCs.

C. Scatter plot comparing fold changes in gene expression induced by TDP-43K263E in iPSCs with those induced by TDP-43 knockdown in ESCs (Modic et al., 2019). DEGs in the study by Modic et al. (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-43K263E in iPSCs with those induced by TDP-43 knockdown in ESCs. Differentially used exons in the study by Modic et al. (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-43K263E iPSC cultures. Middle panel: RNA-seq read coverage in ESC cultures with TDP-43 knockdown. Bottom panel: 3′ end-seq coverage tracks in ESC cultures with TDP-43 knockdown.

H, Schematic illustrating neuron-specific effects of the K263E mutation on the function of TDP-43. In neuronal cell cultures, TDP-43K263E phenocopies TDP-43 knockdown, including changes in gene expression, misregulation of RNA splicing, and aberrant 3′ end processing. On the other hand, in nonneuronal cell types, TDP-43K263E does not exhibit loss-of-function phenotypes.
Figure 1–1. Quality of RNA-seq data
A. RNA integrity number (RIN) of RNA samples subjected to sequencing. B. Average Phred score of pre-filtered sequencing data per read position. C. Filtering result by fastp. D, E. Mapping rate of alignment by HISAT2 (D) and pseudo-alignment by Salmon (E). F. The variance of DEGs in neurons derived from TDP-43<sup>K263E</sup> iPSCs relative to those from wild-type iPSCs.

Figure 2–1. UNC13A splicing
A. Schematic of alternative splicing of UNC13A. CE, cryptic exon. B. Reads that span either exon 20–exon 21 junction, exon 20–CE junction, or CE–exon 21 junction were quantified. C. Percentages of cryptic splicing of UNC13A based on junction spanning read counts. The numbers written above the graphs are the number of samples in which cryptic splicing was detected. D. Schematic of alternative splicing of STMN2. E. Reads that span either exon 1–exon 2 junction or exon 1–CE junction were quantified. F. Percentages of cryptic splicing of STMN2 based on junction spanning read counts. G. Genomic DNA sequencing of risk SNPs in UNC13A intron. H. RT-qPCR confirmed inclusion of CE in UNC13A mRNA in TDP-43<sup>K263E</sup> iPSC-derived neurons (normalized to ACTB; n = 3). I. Percentages of cryptic splicing of UNC13A based on junction spanning read counts in Klim et al. (2019) and Melamed et al. (2019). The numbers written above the graphs are the number of samples in which cryptic splicing was detected. *One sample in Melamed et al. contained no reads spanning exon 20–exon 21 junction, and was excluded for the splicing rate quantification.
This study: log₂FC of gene expression \([K263E/WT]\)  
Klim et al: log₂FC of gene expression \([K263E/WT]\)  
Melamed et al: log₂FC of gene expression \([siTDP/siCTRL]\)  
This study: –log₁₀(adj. \(p\)-value)  
Klim et al: –log₁₀(adj. \(p\)-value)  
Melamed et al: –log₁₀(adj. \(p\)-value)
This study:

Klim et al:
siCTRL
siTDP

Melamed et al:
siCTRL
siTDP

POLDIP3
STMN2
SYT7
PFKP
ELAVL3
DNAJC5
KALRN
ICA1
RNASET2
UNC13B

Spliced intron clusters ranked by adj. p-value

log_{10} FC of exon usage (K263E/WT)

log_{10} FC of exon usage (siTDP/siCTRL)

adj. p = 5e-5

adj. p = 1e-5

adj. p = 1e-8

adj. p = 1e-3
A

**TARDBP**

1 2 3 4 5 6

Protein synthesis

Degradation

C

Exclusive usage of proximal pA site: \( \Psi = 0 \)

Exclusive usage of distal pA site: \( \Psi = 1 \)

D

\[ z \text{-score of } \Psi \]

E

F

\[ \text{chr1:11,018,300-11,032,000} \]

**WT rep1**

**WT rep2**

**WT rep3**

**K263E rep1**

**K263E rep2**

**K263E rep3**

\[ \text{chr1:131,954,406-131,960,319} \]

**WT rep1**

**WT rep2**

**WT rep3**

**K263E rep1**

**K263E rep2**

**K263E rep3**

\[ \text{chr10:53,373,532-53,380,557} \]

**PPP2R2D**

**SMC1A**
**A**

WT  
K263E  

TDP-43  
Hoechst  
TUBB3  
Merge

**B**

WT  

![Graph showing intensity vs. distance](image)

**C**

TDP-43/TubB3 correlation coefficient

WT  
K263E
This study (iPSC):

WT
K263E

Modic et al (ESC):

shCTRL
shTDP

10
100
200
500

Expression

log\(2\) of exon usage [K263E/WT]

log\(2\) of exon usage [shTDP/shCTRL]

0.01
0.10
1.00

iPSC
NPC
Neuron

Fold change [K263E/WT]

TARDBP

Δ

Intron7

PPP2R2D

131,956 kb 131,958 kb 131,960 kb

chr10:131,954,742-131,960,655

WT rep1
WT rep2
WT rep3
K263E rep1
K263E rep2
K263E rep3
shCTRL
shTDP

#reads

0
25
50

RNA-seq

3'end-seq

This study (iPSC)

Modic et al (ESC)

TDP43

knockdown

K263E (FTLD mutation)

• Gene expression change
• RNA missplicing
• Impaired 3' end processing

Neuron

Non-neuronal cell (e.g. iPSC)