

TDP-43 nuclear loss in FTD/ALS causes widespread alternative polyadenylation changes

Received: 15 January 2024

Accepted: 1 July 2025

Published online: 21 October 2025

 Check for updates

Yi Zeng^{1,2}✉, Anastasiia Lovchykova¹, Tetsuya Akiyama¹,
Stephanie L. Rayner^{1,3}, Vidhya Maheswari Jawahar^{4,5}, Chang Liu¹,
Odilia Sianto¹, Caiwei Guo¹, Anna Calliari^{4,5}, Mercedes Prudencio^{4,5},
Dennis W. Dickson^{4,5}, Leonard Petrucelli^{4,5} & Aaron D. Gitler^{1,2,6}✉

In frontotemporal dementia and amyotrophic lateral sclerosis, the RNA-binding protein TDP-43 is depleted from the nucleus of neurons in the brain and spinal cord. A key function of TDP-43 has emerged as a repressor of cryptic exon inclusion during pre-mRNA splicing, but a role for TDP-43 in other RNA-processing events remains unresolved. Here we show that loss of TDP-43 from neuronal nuclei of human brain and disease-causing mutations in TDP-43 are associated with widespread changes in alternative polyadenylation (APA). Using high-resolution polyadenylation site mapping, we comprehensively defined TDP-43-regulated APA events in human stem cell-derived neurons and found that both the strength and position of TDP-43 binding influence polyA site usage. APA events caused by loss of TDP-43 impact expression of disease-relevant genes (for example, *SFPQ*, *NEFL* and *TMEM106B*). These findings provide evidence that, in addition to cryptic exon inclusion, APA changes are a new facet of TDP-43 pathology.

TDP-43 binds to uridine/guanine-rich motifs in RNA transcripts^{1,2} and has a critical role in various aspects of RNA metabolism^{3–5}. Defects in RNA metabolism are considered central to frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) pathogenesis³. However, it is not yet fully understood what genes TDP-43 regulates, what aspect of their RNA metabolism TDP-43 regulates and how their dysregulation promotes neurodegeneration. A major role of TDP-43 has emerged as a repressor of so-called cryptic exons during splicing⁶. Cryptic exons reside in introns of genes and are normally excluded from mature mRNAs. When TDP-43 is dysfunctional (that is, when depleted from the nucleus in FTD/ALS), these cryptic exons are spliced into final mRNAs, often leading to frameshifts, decreased RNA stability or even

the production of new peptide sequences^{7–13}. Notably, some of these cryptic exons are in genes critical for neuronal functions (for example, *STMN2* (refs. 7–9,14)) or genes that harbor disease-associated variants that sensitize them to cryptic splicing upon loss of TDP-43 (for example, *UNC13A*^{10,11}). Cryptic splicing events could serve as powerful biomarkers for TDP-43 dysfunction^{12,13} or even as therapeutic targets^{15,16}.

Besides its well-established role in splicing, TDP-43 also has a critical role in other aspects of RNA processing. Do some of these additional TDP-43-dependent RNA-processing pathways also contribute to disease? One potential disease-relevant RNA-processing pathway is alternative polyadenylation (APA)—a major layer of gene regulation that occurs in >60% of human genes¹⁷. When a gene is transcribed into

¹Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. ²The Phil and Penny Knight Initiative for Brain Resilience, Stanford University, Stanford, CA, USA. ³Motor Neuron Disease Research Centre, Macquarie Medical School, Faculty of Medicine, Health, and Human Sciences, Macquarie University, Sydney, New South Wales, Australia. ⁴Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA. ⁵Neuroscience Graduate Program, Mayo Clinic Graduate School of Biomedical Sciences, Jacksonville, FL, USA. ⁶Chan Zuckerberg Biohub—San Francisco, San Francisco, CA, USA. ✉e-mail: yizeng8@stanford.edu; agitler@stanford.edu

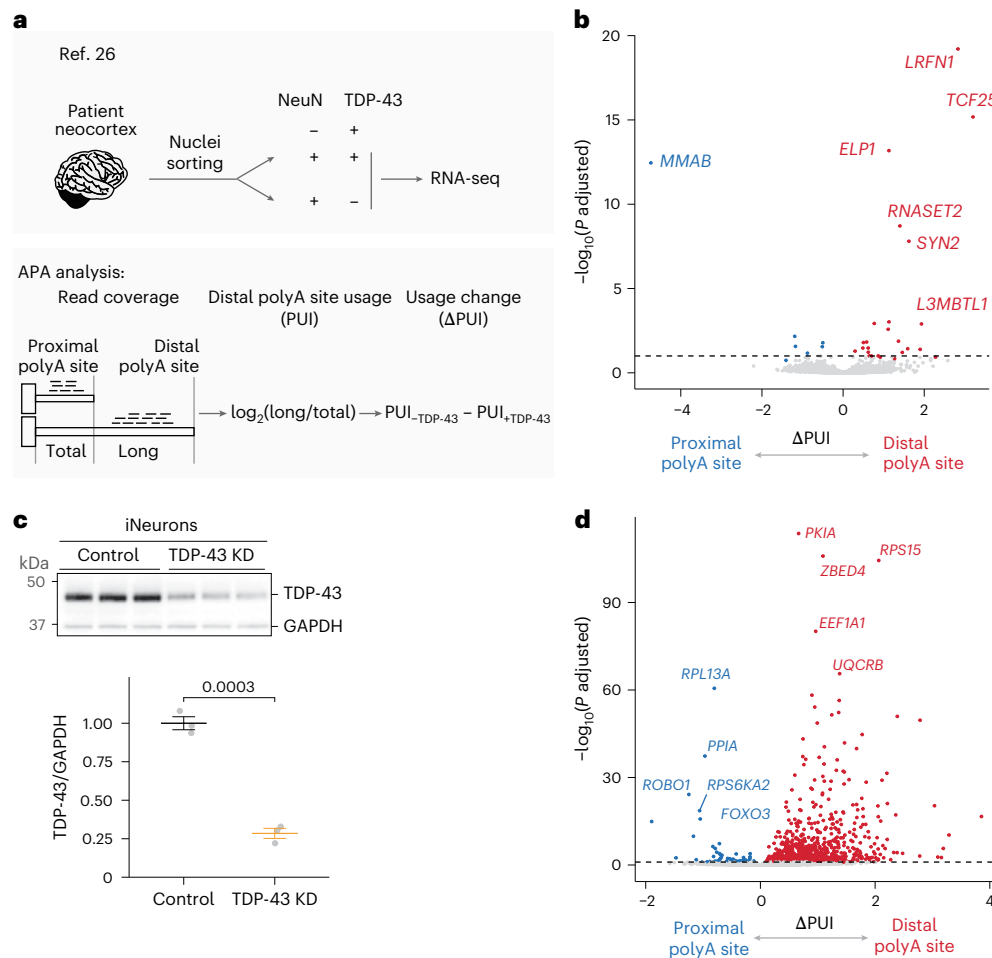


Fig. 1 | Loss of TDP-43 leads to widespread APA changes. **a**, APA analysis scheme. Scheme for sorting neuronal nuclei from frontal cortices of FTD or FTD-ALS patients based on TDP-43 levels for RNA-seq (top)²⁶. Workflow of APA analysis using RNA-seq data (bottom). The Δ PUI is defined as the difference in \log_2 -transformed distal polyA site usage between TDP-43 negative nuclei and TDP-43 positive nuclei. **b**, Loss of TDP-43 is associated with APA changes. APA change (Δ PUI by APALyzer) is plotted on the x axis and the corresponding statistical significance (the adjusted P value is calculated using the Benjamini-Hochberg procedure) on the y axis. Genes with statistically significant APA

change (adjusted $P < 0.1$) are labeled in red for favoring distal polyA sites and in blue for favoring proximal polyA sites. Dashed horizontal line marks the adjusted $P = 0.1$. **c**, Western blot shows efficient TDP-43 KD after 7 days. Quantitation of protein levels is presented as mean \pm s.e.m.; the P value was calculated using one-sided Student's t -test; control, samples treated with scrambled shRNA ($n = 3$); TDP-43 KD, samples treated with TDP-43 shRNA ($n = 3$). **d**, Volcano plot shows widespread APA changes upon TDP-43 KD in iNeurons; thresholds for significant changes are $|\Delta$ PUI| > 0.10 , calculated by APALyzer as in **b**, and adjusted $P < 0.05$, calculated as in **b**.

mRNA, it is cleaved and polyadenylated, which functions to stabilize the mRNA, facilitate its nuclear export and regulate its translation¹⁸. If alternative polyadenylation (polyA) sites are used, it could produce mRNA isoforms that have different 3' untranslated region (UTR) lengths, impacting RNA and protein levels, subcellular localization or even protein functions^{17,19}. If APA occurs prematurely, it can truncate mRNAs, reducing full-length protein levels^{17,19}. Previous genome-wide TDP-43 binding studies revealed that TDP-43 binding is enriched not only in introns of genes but also in 3' UTRs^{1,2}, suggesting a role in regulating polyadenylation. Indeed, TDP-43 knockdown (KD) or disease-associated TDP-43 mutations affect polyadenylation²⁰⁻²², including its own polyadenylation^{2,23} and premature polyadenylation in *STMN2* (refs. 7,8). Notably, widespread APA changes have been observed in ALS patient samples using bulk RNA sequencing (RNA-seq)²⁴ and single-nucleus RNA-seq²⁵, although it is unclear whether these changes are directly owing to TDP-43 dysfunction.

Here we show that loss of TDP-43 from neuronal nuclei in postmortem brain tissue of FTD/ALS patients is associated with APA changes. Using high-resolution polyadenylation mapping in human stem

cell-derived neurons, we comprehensively defined TDP-43-regulated APA events and discovered that TDP-43 loss caused widespread APA changes, impacting expression of disease-relevant genes, such as *NEFL*, *SFPQ* and *TMEM106B*. Notably, the strength and position of TDP-43 binding, along with the relative strengths of polyA sites, influence the outcome of APA changes upon TDP-43 dysfunction. Together, these findings provide evidence for a crucial role of TDP-43-mediated APA in maintaining neuronal health and its dysregulation as a key contributor to neurodegenerative disease.

Results

TDP-43 loss is linked to APA changes

To test the hypothesis that APA caused by TDP-43 dysfunction contributes to FTD/ALS, we first performed APA analysis in an RNA-seq dataset²⁶, in which neuronal nuclei with or without TDP-43 were sorted from FTD/ALS postmortem brain samples for RNA-seq (Fig. 1a, top). We recently re-analyzed this dataset to discover TDP-43-dependent cryptic splicing events¹¹. Here we 're-re-analyzed' it to search for APA events (Fig. 1a, bottom). We used two different APA analysis programs

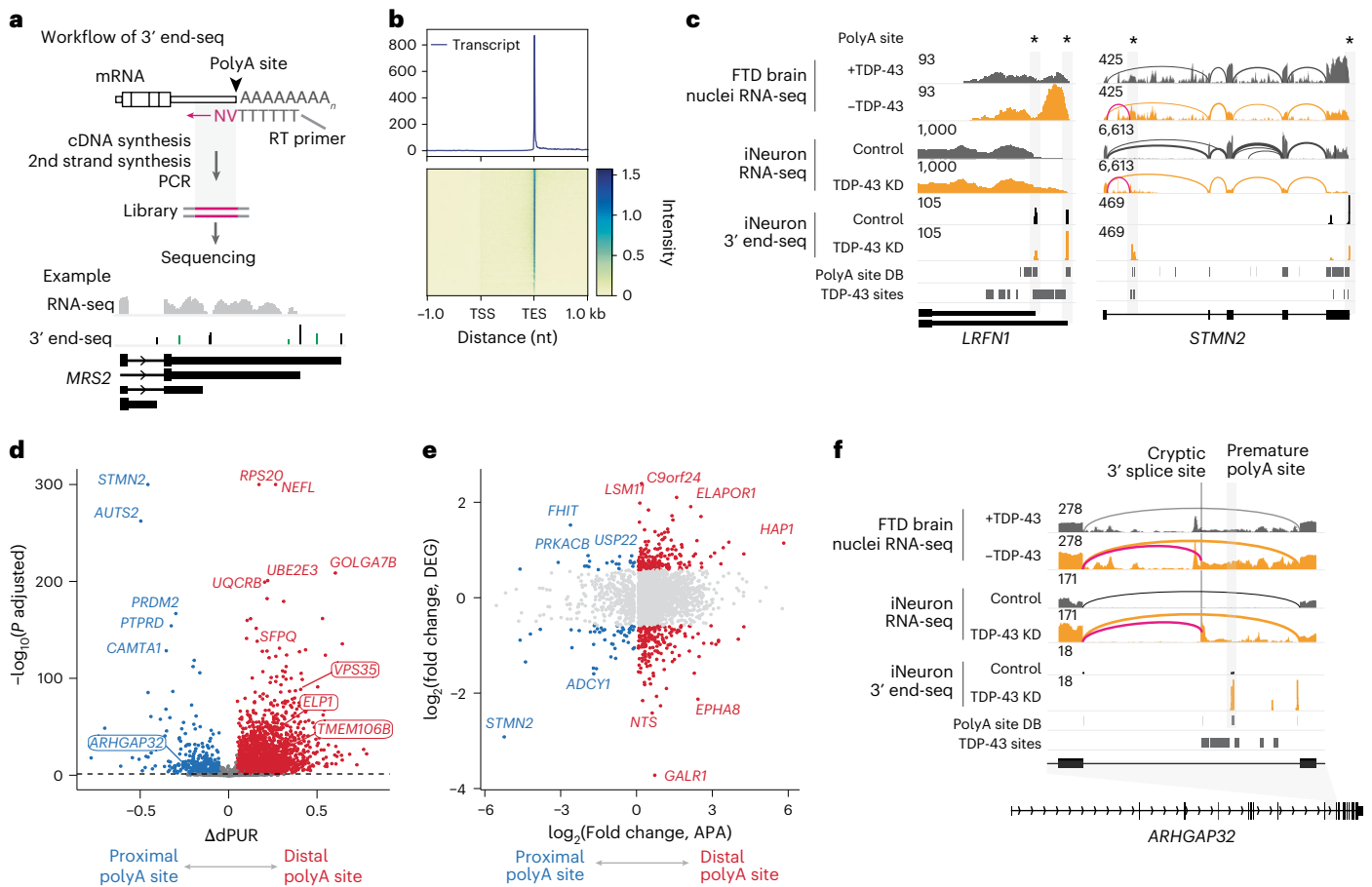


Fig. 2 | High-resolution mapping of TDP-43-dependent APA using 3' end-seq. **a**, Workflow of 3' end-seq (top). Example of 3' end-seq read coverage, indicative of polyA sites, at the 3' UTR of *MRS2* (bottom). Top track marks RNA-seq read coverage, middle track marks read coverage of 3' end-seq with new sites marked in green and bottom track marks gene structure. **b**, Coverage of filtered read ends across transcripts. TSS, transcription start site; TES, transcription end site. **c**, Example of TDP-43-regulated APA changes detected by 3' end-seq. The direction of *LRFNI* gene is flipped for illustrative purpose (left). Example of premature polyadenylation captured by 3' end-seq (right). Sashimi plots in the top four tracks illustrate splicing; the pink curved line marks the usage of a cryptic 3' splice site. The 'polyA DB' track marks annotated polyA sites³¹ and the 'TDP-43 sites' track marks the observed TDP-43 binding³⁴. Gene structure is shown at the

bottom. The asterisk marks the positions of observed polyA site. **d**, Volcano plot showing 3' end-seq captured widespread APA changes upon TDP-43 KD. The APA change (ΔdPUR) is defined by the difference in the distal polyA site usage between TDP-43 KD and control samples. Thresholds for significant changes are $|\Delta\text{dPUR}| > 0.05$ and adjusted $P < 0.05$. The adjusted P value is calculated using the Benjamini–Hochberg procedure. **e**, Scatter plot with APA changes, represented by distal polyA site usage, on the x axis, and RNA level changes on the y axis, illustrating that APA changes could alter RNA levels. Genes with significant APA changes are plotted; genes with significant RNA-level changes are labeled in red when favoring distal polyA sites or blue when favoring proximal polyA sites. **f**, TDP-43 KD activated cryptic splicing and premature polyadenylation in intron 12 of *ARHGAP32*. Details are shown in **c**.

and identified 41 APA changes using APALyzer²⁷ and 62 APA changes using QAPA²⁸ (Fig. 1b, Supplementary Fig. 1a and Supplementary Table 1; adjusted $P < 0.1$). Only two genes with substantial APA changes were common between both programs—*LRFNI* and *MARK3* (Fig. 2c and Supplementary Fig. 2l), likely because of different polyA databases used by each program and the low resolution of RNA-seq for APA analysis. We thus considered APA changes identified by each program, which revealed APA events in genes critical for neuronal function, such as *LRFNI* and *SYN2*. APA changes in 34 genes are associated with gene expression changes (Supplementary Fig. 1b), likely impacting their corresponding protein levels (see below).

To test whether TDP-43 directly regulates APA, we knocked down TDP-43 levels in cortical neurons differentiated from human stem cells (iNeurons) and performed RNA-seq for APA analysis (Fig. 1c and Supplementary Fig. 1c,d). TDP-43 KD caused widespread APA changes (Fig. 1d and Supplementary Table 2; $|\Delta\text{PUI}| > 0.1$ and adjusted $P < 0.05$ from APALyzer), including 24 changes, that were also observed in FTD/ALS postmortem brain samples (Fig. 1b and Supplementary Fig. 1a). Applying the same analysis to published RNA-seq datasets, we also

observed APA changes in human-induced pluripotent stem cell-derived motor neurons (iMNs) upon TDP-43 KD⁷ (Supplementary Fig. 1e), in iNeurons carrying a pathogenic mutation (TDP-43-K263E), which disrupts RNA binding²⁹ (Supplementary Fig. 1f), and in patient-derived iMNs carrying a different pathogenic mutation (TDP-43-M337V), which disrupts TDP-43 intermolecular interactions²¹ (Supplementary Fig. 1g). Although technical differences among these studies preclude direct comparisons, the presence of widespread APA changes across different studies strongly suggests that TDP-43 regulates APA in neurons and that TDP-43 loss of function or ALS-linked mutations result in widespread APA changes owing to direct and indirect effects of TDP-43 dysfunction.

Comprehensive high-resolution mapping of TDP-43-dependent APA

To comprehensively map TDP-43-dependent APA events and define how TDP-43 dysfunction contributes to neurodegeneration through APA, we sought a more sensitive assay than RNA-seq for APA analysis, because RNA-seq cannot map new polyadenylation events, identify actual polyA sites or capture 'internal' premature polyadenylation events (that is,

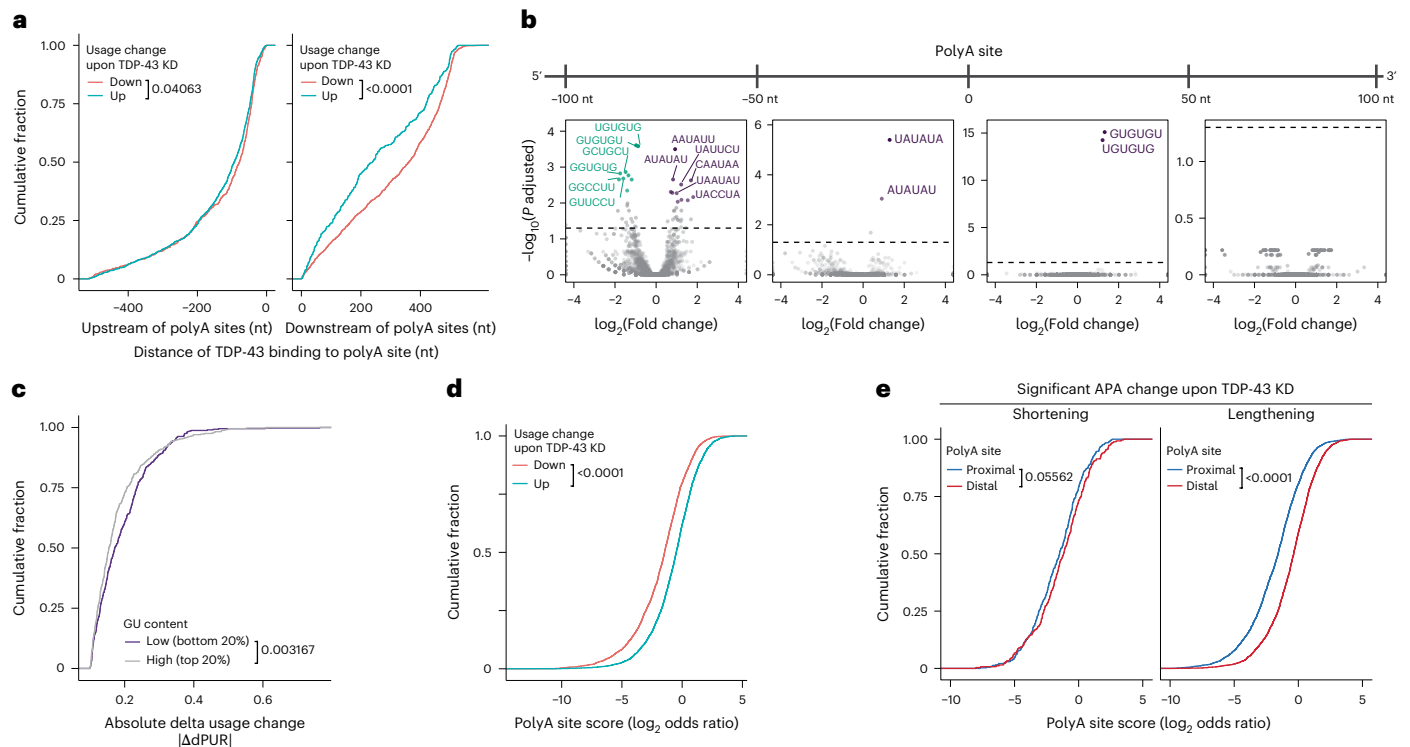


Fig. 3 | The position and strength of TDP-43 binding influence polyA site usage. **a**, Cumulative plots showing the location of TDP-43 binding relative to TDP-43-regulated polyA sites. **b**, Volcano plots showing the enrichment of hexamers in different regions surrounding polyA sites. **c**, Cumulative plots show that TDP-43 KD caused a larger absolute usage change in polyA sites with weak TDP-43 binding (low GU content) than sites with strong TDP-43 binding (high GU content). **d**, Cumulative plots show that for TDP-43-regulated APA sites, polyA

sites with increased usage are stronger than polyA sites with decreased usage. **e**, Cumulative plots show that for genes with significant 3' UTR lengthening upon TDP-43 KD, their distal polyA sites are stronger than their proximal ones. PolyA site scores (in **d** and **e**) were calculated using Aparent2, expressed as log odds ratio, and reflect polyA site strengths. P values in **a** and **c–e** were calculated using two-sided Mann–Whitney U test, and adjusted P values in **b** were calculated using the Benjamini–Hochberg procedure.

ones that truncate the RNA/protein). We used a specialized transcriptomic method, 3' end-seq, to map polyA sites with single-nucleotide resolution in iNeurons with or without TDP-43 KD (Fig. 2a). We used stringent filtering criteria to remove polyA sites likely derived from mis-priming events during cDNA synthesis, sites with no known polyA signals or sites with low read coverage (see Methods for details), resulting in high-quality polyA sites with read ends that align exactly at the 3' ends of annotated transcripts (Fig. 2b). In total, we identified 60,369 polyA sites with a known upstream polyA signal, of which 7,975 polyA sites are new, when compared with a compendium of polyA site annotations^{30,31}. Both annotated and new polyA sites harbored known polyA signals located ~ 20 nts upstream (Supplementary Fig. 2a,b). Consistent with previous findings^{2,20}, the majority of polyA sites we identified were in 3' UTRs, though the new polyA sites had more diverse distributions within transcripts (Supplementary Fig. 2c), consistent with the idea that they are normally repressed by TDP-43 under normal conditions. Furthermore, 3' end-seq confirmed that loss of TDP-43 lengthened the 3' UTR of *LRFNI* (Fig. 2c, left), which we also observed in our analysis of FTD/ALS postmortem brain samples (Fig. 1b and Supplementary Fig. 1a). Notably, 3' end-seq also captured premature polyadenylation in *STMN2*, which was previously identified but not apparent in standard RNA-seq data (Fig. 2c, right), demonstrating the power of 3' end-seq for studying premature polyadenylation, the most detrimental form of APA, because it truncates the RNA/protein.

In total, TDP-43 KD altered the usage of 7,304 polyA sites ($|\Delta\text{dPUR}| > 0.05$ and adjusted $P < 0.05$) and caused APA changes in 3,206 genes (Fig. 2d and Supplementary Tables 3 and 4). Like recent findings in ALS samples²⁴ and FTD samples (Fig. 1b and Supplementary Fig. 1a), as well as in our RNA-seq data from iNeurons (Fig. 1d), the majority of

genes with APA changes (2,752 genes) detected by 3' end-seq lengthened their RNA transcripts; 433 APA events were associated with at least a 1.5-fold change in RNA level (Fig. 2e), suggesting that APA changes could alter gene expression. Notably, we also observed several of these APA changes in FTD/ALS postmortem brain samples, and they were in genes connected to ALS or FTD (Supplementary Table 5).

In addition, 3' end-seq also empowered the discovery of 'cryptic' polyA sites (ones not used under normal conditions but revealed upon TDP-43 KD). TDP-43 KD activated 404 cryptic polyA sites in 372 genes (Supplementary Table 3); 65 cryptic polyA sites occurred downstream of a gene's annotated 3' end-seq, such as the FTD risk factor *RFNG*³² (Supplementary Fig. 2d), *ELP6* (Supplementary Fig. 3e) and *SIX3*, *TLX1* and *ELK1* (Supplementary Fig. 2e–g; ref. 33). In addition, 149 cryptic events induced premature polyadenylation in genes such as *TNIP1*, *EGFR* and *GSTO2* (Supplementary Fig. 2h–j). Like coordinated activation of splicing and polyadenylation in intron 1 of *STMN2*, we observed apparent coordination of these two processes in 13 additional genes (Supplementary Table 6), such as *ARHGAP32* (Fig. 2f; ref. 33). Cryptic splicing of *ARHGAP32* was recently reported¹², and we also confirmed this by qRT–PCR (Supplementary Fig. 2k) and in RNA-seq data from FTD/ALS postmortem brain samples (Fig. 2f). These results suggest a potential coupling between cryptic splicing and the activation of cryptic polyadenylation sites.

The position and strength of TDP-43 binding influence APA

By cross-referencing TDP-43-regulated APA sites with a curated genome-wide TDP-43-binding site dataset³⁴, we found that $\sim 70\%$ of genes with these APA events had at least one TDP-43-binding site. Interestingly, TDP-43-binding sites located downstream of polyA sites were

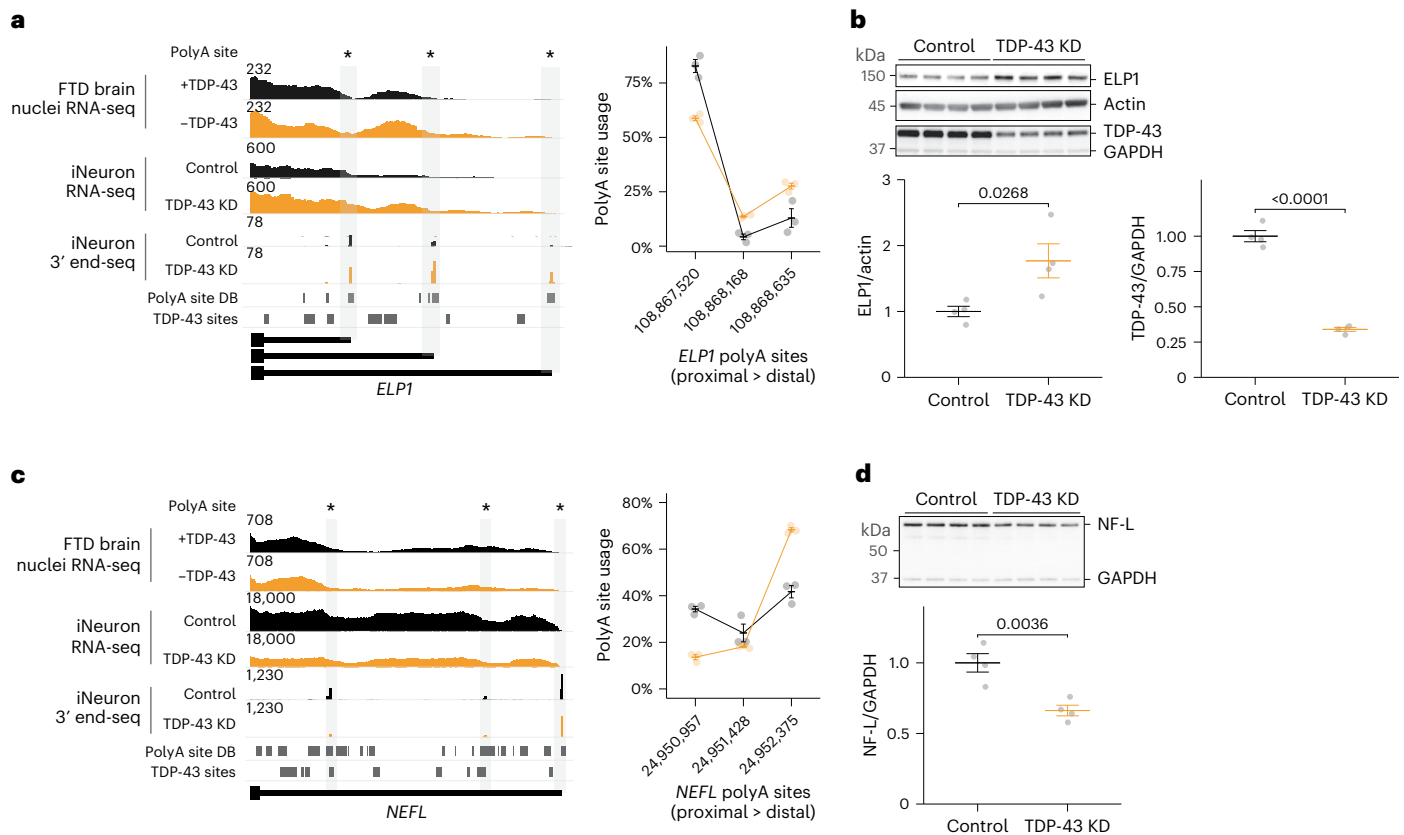


Fig. 4 | Loss of TDP-43 causes APA changes in *ELPI* and *NEFL*, impacting their protein levels. **a, TDP-43 KD lengthened the 3' UTR of *ELPI*. **b**, TDP-43 KD increased *ELPI* protein levels. Quantitation of protein levels ($n = 4$ per condition) is presented as mean \pm s.e.m.; P values were calculated using one-sided Student's**

t -test. **c**, TDP-43 KD downregulates the usage of the proximal polyA site of *NEFL*, leading to the increased usage of the most distal polyA site. The asterisk in **a** and **c** marks the positions of observed polyA site. **d**, TDP-43 KD reduced *NF-L* protein levels. Quantitation of protein levels ($n = 4$ per condition) as in **a**.

found closer to APA sites with increased usage, and farther from those with decreased usage, upon TDP-43 KD (Fig. 3a). Consistent with this observation, strong TDP-43 binding was enriched 50–100 nts upstream of APA sites with reduced usage, whereas two of the strongest TDP-43 binding motifs (GUGUGU and UGUGUG) were enriched 0–50 nts downstream of APA sites with increased usage (Fig. 3b). Notably, these two GU-repeat containing hexamers were also binding sites for cleavage stimulation factor 2 (CstF2)¹⁹, which binds immediately downstream of a polyA site to facilitate its cleavage and polyadenylation. Thus, TDP-43 might block CstF2's binding to inhibit polyA site usage, and when TDP-43 is absent, CstF2 can bind to the region and promote polyA site usage.

Using GU content to estimate the strength of TDP-43 binding located downstream of APA sites, we found that TDP-43 KD caused a larger usage change in polyA sites with weak TDP-43 binding than sites with strong sites (Fig. 3c and Supplementary Fig. 2o), suggesting that, when reduced, TDP-43 likely dissociates from weak sites first, then from stronger ones.

PolyA site strength influences TDP-43-dependent APA

To determine whether polyA site strength influences TDP-43-regulated APA, we estimated the probability of cleavage and polyadenylation for each identified polyA site using a deep residual neural network-based model, Aparent2 (ref. 35), that accurately predicts polyA site usage and strength. Interestingly, we found that polyA sites with decreased usage upon TDP-43 KD are weaker than those with increased usage upon TDP-43 KD (Fig. 3d). Whereas genes with substantial 3' UTR shortening show no difference in polyA strength between proximal and distal polyA sites, genes with substantial 3' UTR lengthening have stronger distal sites compared with proximal ones (Fig. 3e). Given the cotranscriptional

nature of cleavage and polyadenylation, our findings suggest that in the case of 3' UTR lengthening upon TDP-43 KD, a stronger distal polyA site might compensate for its positional disadvantage over the proximal polyA site during transcription and therefore gets activated upon TDP-43 KD.

TDP-43 loss induces APA changes of disease-associated genes

TDP-43 depletion increased the usage of two distal polyA sites in *ELPI* that lengthened its 3' UTR (Fig. 4a). We confirmed such lengthening by qRT-PCR (Supplementary Fig. 3a) and longer *ELPI* 3' UTR is found in multiple published datasets that have TDP-43 dysfunction (Supplementary Fig. 3b; ref. 36). *ELPI*, also called *IKBKAP*, encodes a subunit of the elongator complex, which mediates the methylation of the wobble base in tRNAs, and has been functionally and genetically linked to ALS^{37–39}. In addition to *ELPI*, we observed APA changes in two other subunits of the elongator complex (*ELP3* and *ELP6*) upon TDP-43 KD (Supplementary Fig. 3c–e). These APA changes are associated with protein levels; TDP-43 KD led to increased protein levels of *ELPI* and *ELP3* (Fig. 4b and Supplementary Fig. 3d). Increased *ELPI* protein levels may be due to increased *ELPI* RNA levels (Supplementary Fig. 3c). These results suggest a possible contribution of abnormal tRNA metabolism to FTD/ALS.

Previous studies found that TDP-43 directly binds to the 3' UTR of *NEFL* to stabilize its RNA levels⁴⁰. *NEFL* encodes neurofilament light (NF-L) chain, which has emerged as a sensitive prognostic biomarker for diverse neurodegenerative diseases⁴¹, including FTD/ALS⁴². TDP-43 KD not only reduced *NEFL* RNA levels (Supplementary Fig. 3f), consistent with previous findings⁴³, but also further shifted the apparent polyA site usage from proximal to distal (Fig. 4c), which we confirmed by

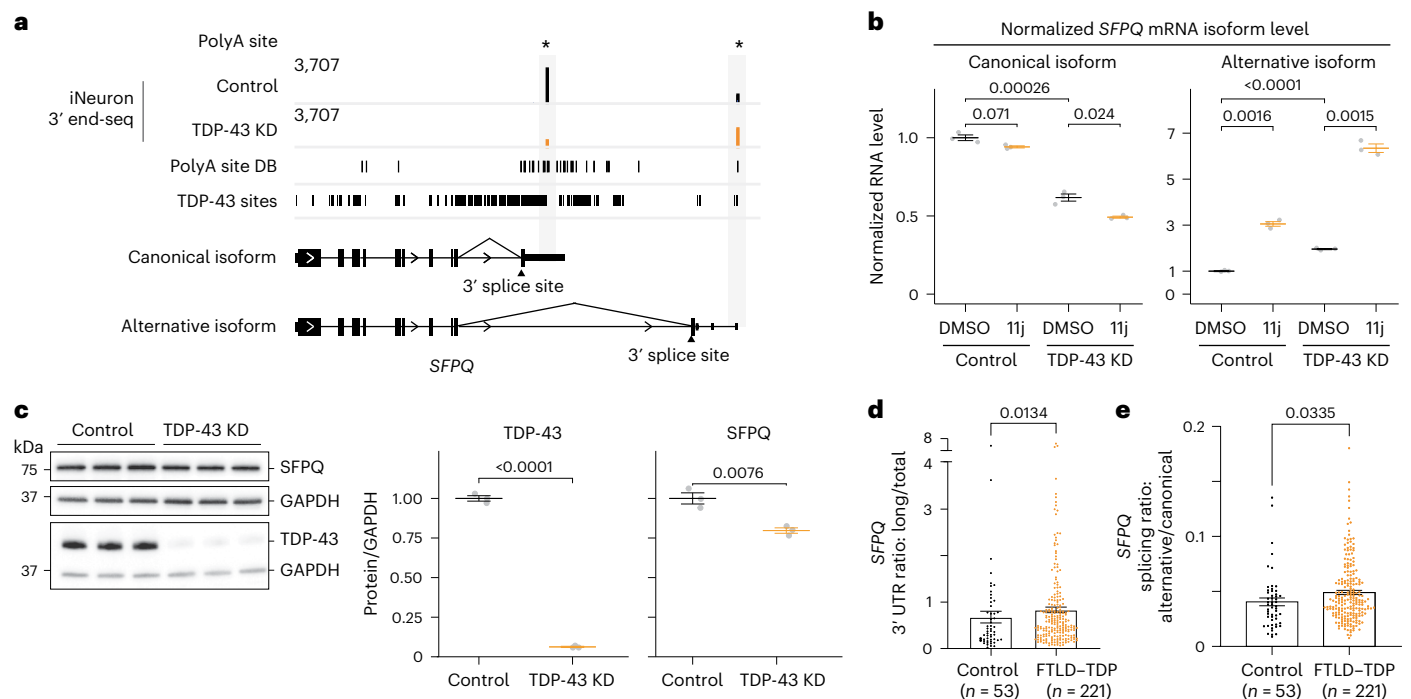


Fig. 5 | Loss of TDP-43 promotes the usage of a distal polyA site and a downstream 3' splice site in *SFPQ*, resulting in an NMD-sensitive mRNA isoform. **a**, TDP-43 KD shifted the polyA site usage from proximal to distal in *SFPQ*. The gene structure illustrates the following two mRNA isoforms: the canonical mRNA isoform that uses a proximal polyA site and an upstream 3' splice site of intron 9; the alternative mRNA isoform that uses a distal polyA site and a downstream 3' splice site of intron 9. The presence of additional exons downstream of the termination codon in the alternative mRNA isoform suggests it to be a potential NMD target. **b**, Inhibiting NMD with a small molecule (11j at

0.1 μ m) stabilized the alternative *SFPQ* mRNA isoform in both control and TDP-43 KD conditions ($n = 4$ per condition). Data are presented as mean \pm s.e.m.; P values were calculated using one-sided Student's t -test. **c**, Western blot shows that TDP-43 KD reduced *SFPQ* protein levels. Quantitation of protein levels ($n = 3$ per condition) is presented as mean \pm s.e.m.; P values were calculated using one-sided Student's t -test. **d, e**, Levels of the long 3' UTR (**d**) and the alternative splicing isoform (**e**) of *SFPQ* are substantially increased in the frontal cortices of FTLD-TDP patients ($n = 221$) compared with controls ($n = 53$). Data are presented as mean \pm s.e.m.; P values were calculated using two-tailed Mann-Whitney test.

qRT-PCR (Supplementary Fig. 3g), and reduced NF-L protein levels (Fig. 4d). These observations suggest that, together with *NEFL* 3' UTR-targeting microRNAs⁴⁴, loss of TDP-43-induced APA changes might further reduce NF-L levels.

TDP-43 KD also shifted polyA site usage from proximal to distal in another FTD/ALS-linked gene, *SFPQ* (Fig. 5a and Supplementary Fig. 3h), which encodes a ubiquitously expressed RNA-binding protein that has key roles in RNA metabolism^{45,46}, and is depleted from the nucleus in sporadic ALS spinal cord⁴⁷, losing its interaction with the FTD/ALS gene *FUS*, in neuronal nuclei in FTLD-TDP^{48,49}. The gene structure of *SFPQ* suggests that, when the distal polyA site in *SFPQ* is used, a downstream 3' splice site of intron 9 might be used (Fig. 5a and Supplementary Fig. 3h). Indeed, we confirmed that TDP-43 KD promoted the usage of both the distal polyA site (Supplementary Fig. 3i) and the alternative 3' splice site (Supplementary Fig. 3j).

Intriguingly, this alternative *SFPQ* mRNA isoform seems to be a nonsense-mediated decay (NMD) target (Fig. 5a), as there are additional exons downstream of the termination codon. Indeed, inhibiting NMD increased the level of this alternative *SFPQ* mRNA isoform (Fig. 5b and Supplementary Fig. 3k), and this alternative isoform is less stable than the canonical isoform (Supplementary Fig. 3l). Notably, TDP-43 KD also reduced *SFPQ* protein levels (Fig. 5c). Together, these results provide evidence that TDP-43 KD promotes the production of an *SFPQ* alternative mRNA isoform that is an NMD target, leading to reduced *SFPQ* protein levels.

To validate the functional significance of the APA change in *SFPQ*, we analyzed the usage of the distal polyA site and the alternative 3' splice site in a series of 274 frontal cortex brain samples from the Mayo Clinic Brain Bank using qRT-PCR. We found a substantial increase in

the level of the longer *SFPQ* 3' UTR and the usage of the alternative 3' splice site in patients with FTLD-TDP compared with control samples (Fig. 5d,e), indicating that this *SFPQ* APA event occurs in human disease.

FTD risk gene *TMEM106B* is a TDP-43-regulated APA target

The sensitivity of 3' end-seq further revealed that TDP-43 KD lengthened *TMEM106B*'s 3' UTR (Fig. 6a), which we also confirmed by qRT-PCR (Supplementary Fig. 3m). *TMEM106B* is a top genetic risk factor that emerged in a genome-wide association study for FTLD-TDP⁵⁰. C-terminal fragments of *TMEM106B* were recently found to form amyloid fibrils in the brains of older individuals and patients with neurodegenerative disorders⁵¹⁻⁵⁴. By targeted analysis of RNA-seq read coverage across the *TMEM106B* 3' UTR, we confirmed that loss of TDP-43 is associated with a longer *TMEM106B* 3' UTR in FTD/ALS postmortem brain samples (Fig. 6b). We also found a substantial increase in longer *TMEM106B* 3' UTRs in the frontal cortices of patients with FTLD-TDP compared with healthy controls (Fig. 6c), indicating that the increased 3' UTR length of *TMEM106B* might be functionally relevant in FTD/ALS.

Because *TMEM106B* protein levels are altered in disease^{50,55}, we next examined whether TDP-43 KD affected *TMEM106B* protein levels. Using a condition that has been shown previously to preserve *TMEM106B* dimers on SDS-PAGE⁵⁵ (Supplementary Fig. 3n), we found that TDP-43 KD did not affect levels of *TMEM106B* monomers (~37 kDa) but reduced its dimer levels (Fig. 6d). A recent proteomics study also detected decreased *TMEM106B* levels caused by TDP-43 KD¹². TDP-43 KD did not substantially affect total *TMEM106B* RNA levels (Supplementary Fig. 3o) but increased the RNA stability of the longer 3' UTR isoform (Supplementary Fig. 3p), suggesting that the longer 3' UTR might affect *TMEM106B* protein levels through translation. To test this hypothesis,

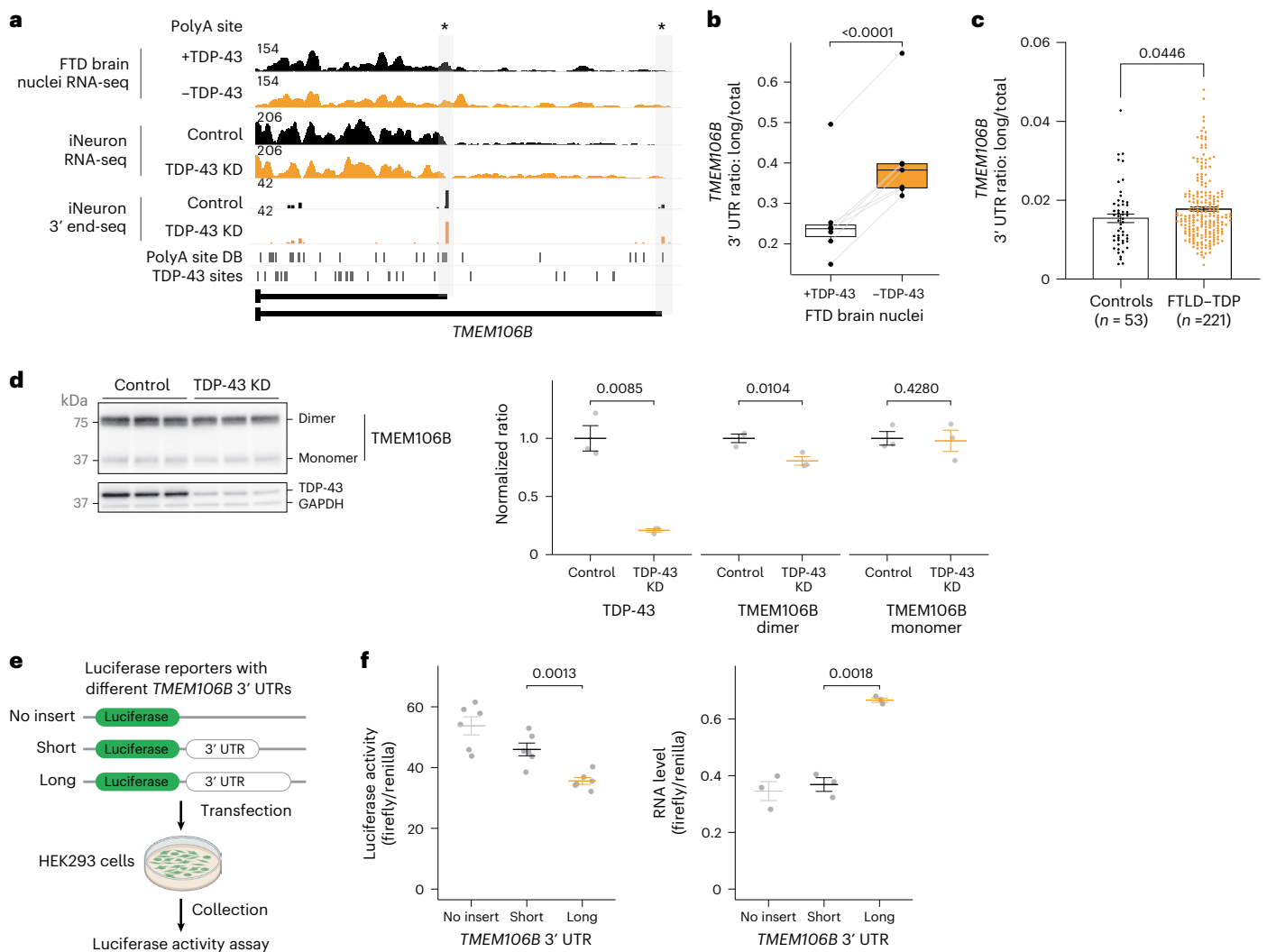


Fig. 6 | FTD risk gene *TMEM106B* is a TDP-43-regulated APA target. **a**, TDP-43 KD lengthened the 3' UTR of *TMEM106B*. **b**, Targeted APA analysis of *TMEM106B* using RNA-seq from FTD/ALS brain samples, shown in box plots with centerline, median; box limits, upper and lower quartiles; and whiskers, minimum and maximum. **c**, Level of the long *TMEM106B* 3' UTR is substantially increased in the frontal cortices of FTLD-TDP patients ($n = 221$) compared with controls ($n = 53$). Data are presented as mean \pm s.e.m.; P values were calculated using two-tailed Mann-Whitney test. **d**, Western blot shows reduced *TMEM106B* dimer levels after 12-day TDP-43 KD. Quantitation of protein levels ($n = 3$ per condition) is

presented as mean \pm s.e.m.; P values were calculated using one-sided Student's t -test. **e**, The schematic of the luciferase activity assay for *TMEM106B* 3' UTRs. **f**, The presence of the long *TMEM106B* 3' UTR reduces translation efficiency. Dot plots show that the longer *TMEM106B* 3' UTR led to lower luciferase activity (left) and higher RNA levels (right). No insert, a dual-luciferase reporter without *TMEM106B* 3' UTRs; short, a dual-luciferase reporter with the short *TMEM106B* 3' UTR; long, a dual-luciferase reporter with the long *TMEM106B* 3' UTR. Data are presented as mean \pm s.e.m.; P values were calculated using one-sided Student's t -test. Panel e created with [BioRender.com](https://www.biorender.com).

we cloned short and long versions of the *TMEM106B* 3' UTR into a dual-luciferase reporter (Fig. 6e). We mutated the proximal polyA site in the long 3' UTR-containing reporter to prevent its usage and confirmed the production of the intended long 3' UTR (Supplementary Fig. 3q). The long 3' UTR-containing reporter had substantially lower luciferase activity and higher RNA levels than the short 3' UTR-containing reporter (Fig. 6f), providing evidence that the loss of TDP-43-induced longer *TMEM106B* 3' UTR reduces translation efficiency. *TMEM106B* dimer levels formed from the construct lacking a 3' UTR were not sensitive to TDP-43 KD (Supplementary Fig. 3r), suggesting that TDP-43 influences *TMEM106B* dimer levels through its 3' UTR.

Discussion

In the present study, we show that TDP-43 dysfunction, a hallmark of FTD/ALS, leads to widespread APA changes in disease-relevant neuronal cell lines and postmortem brain samples of FTD/ALS (Figs. 1b,d and 2d and Supplementary Fig. 1a,e-g). Furthermore, the

interplay between the position and strength of TDP-43 binding and intrinsic polyA site strength influences TDP-43-regulated APA (Fig. 3). Notably, TDP-43 loss-induced APA changes are found in genes critical for neuronal function and disease-associated genes (Figs. 4-6 and Supplementary Fig. 3), impacting their expression (for example, *ELP1*, *NEFL*, *SFPQ* and *TMEM106B*). Together, these findings establish that APA dysregulation is a new pathogenic mechanism in TDP-43 proteinopathies.

Due to the constraints of RNA-seq for detecting APA, we applied 3' end-seq to comprehensively map neuronal polyadenylation on a transcriptomic scale with base-pair resolution and discovered thousands of TDP-43-dependent APA events, including cryptic polyA sites. TDP-43 loss-induced APA changes substantially altered the transcriptome (Fig. 2d,f and Supplementary Fig. 2d-n), primarily through 3' UTR lengthening, but also through 3' UTR shortening, alternative last exon usage and premature polyadenylation. These findings underscore a critical role for TDP-43 in regulating APA.

By examining TDP-43 binding characteristics, we provide evidence that TDP-43-binding sites closer to the polyA generally result in repression of polyA site usage (Fig. 3a,b), consistent with previous findings²⁰. Furthermore, polyA sites with weaker TDP-43 binding showed greater usage changes upon TDP-43 loss (Fig. 3c and Supplementary Fig. 2o), supporting the idea that reducing TDP-43 levels, such as in disease, likely causes it to dissociate first from weak binding sites before the strong ones. Intrinsic polyA site strength also substantially influences these APA events (Fig. 3d,e); stronger distal polyA sites could overcome their positional disadvantages over proximal polyA sites to lengthen 3' UTRs upon TDP-43 loss. The impact of TDP-43 binding, its strength and polyA site strength on TDP-43-regulated APA provides a basis for future biochemical characterization and a toehold to prioritize antisense oligonucleotide-based therapeutic strategies to target specific APA events akin to current approaches underway to target cryptic splicing events¹⁵.

Many of the APA changes we identified were in genes with established links to FTD/ALS and other neurodegenerative diseases, often resulting in altered protein expression. For example, TDP-43 KD led to APA changes and increased protein levels of *ELP1* (Fig. 4a,b), a subunit of the elongator complex involved in tRNA modification. Because abnormal tRNA metabolism affects translation and a recent finding of reduced aminoacylation of tRNA^{Phe} in FTD/ALS⁵⁶, our observations suggest that defects in tRNA metabolism might be an important mechanism contributing to FTD/ALS. Conversely, upon TDP-43 loss, *NEFL* exhibited a shift to distal polyA site usage accompanied by its reduced RNA and protein levels (Fig. 4c,d), adding complexity to its use as a biomarker in TDP-43 proteinopathies. For another FTD/ALS-linked gene, *SFPQ*, TDP-43 loss promoted the usage of a distal polyA site and a downstream 3' splice site (Fig. 5a,d,e and Supplementary Fig. 3i,k), resulting in an alternative mRNA isoform targeted by NMD (Fig. 5b and Supplementary Fig. 3k), leading to reduced *SFPQ* protein (Fig. 5c); such APA and splicing change in *SFPQ* were validated in FTLD-TDP patient brains (Fig. 5d,e).

Notably, TDP-43 loss lengthened the 3' UTR of *TMEM106B*, a top genetic risk factor for FTLD-TDP (Fig. 3a–c). Intriguingly, this 3' UTR lengthening correlated with reduced *TMEM106B* dimer levels (Fig. 3d). But how could this APA event reduce the formation of *TMEM106B* dimers? *TMEM106B* APA might impact dimer levels by influencing the subcellular localization of *TMEM106B* mRNA and its ability to serve as a scaffold for protein–protein interactions, as has been shown for other APA events¹⁷. Recently, an *Alu* element insertion in the *TMEM106B* 3' UTR was found in strong linkage disequilibrium with the top FTLD-TDP risk allele at this locus^{57–59} and the risk haplotype of *TMEM106B* was associated with reduced *TMEM106B* protein dimer levels and increased fibril levels⁶⁰. Given the proximity of these variants to *TMEM106B*'s distal polyA site, future work will explore whether and how these disease-associated genetic variants impact APA and their association with *TMEM106B* dimer levels.

In summary, our findings, together with independent studies, reveal widespread APA changes associated with TDP-43 dysfunction in FTD/ALS^{33,36}. Despite using different approaches, each of the three studies observed common APA changes, underscoring the importance of APA changes during neurodegeneration. Our application of 3' end-seq complements and extends RNA-seq-based APA analysis because it enables de novo identification of polyadenylation events with base-pair resolution, revealing complex and highly sensitive APA changes (for example, *NEFL*, *SFPQ* and *TMEM106B*). Together, these findings and the related studies provide evidence that TDP-43 pathology contributes to disease pathogenesis through not only cryptic splicing but now also changes in APA. Future research should focus on the downstream functional consequences of these specific APA alterations in disease-relevant genes, examine APA changes in other cell types or diseases with TDP-43 pathology and explore whether targeted correction of

these APA events, perhaps using antisense oligonucleotides, could offer therapeutic benefits.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-025-02049-3>.

References

1. Tollervey, J. R. et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
2. Polymenidou, M. et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).
3. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
4. Klim, J. R., Pintacuda, G., Nash, L. A., Juan, I. G. S. & Eggan, K. Connecting TDP-43 pathology with neuropathy. *Trends Neurosci.* **44**, 424–440 (2021).
5. Birsá, N., Bentham, M. P. & Fratta, P. Cytoplasmic functions of TDP-43 and FUS and their role in ALS. *Semin. Cell Dev. Biol.* **99**, 193–201 (2020).
6. Ling, J. P., Pletnikova, O., Troncoso, J. C. & Wong, P. C. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* **349**, 650–655 (2015).
7. Klim, J. R. et al. ALS-implicated protein TDP-43 sustains levels of *STMN2*, a mediator of motor neuron growth and repair. *Nat. Neurosci.* **22**, 167–179 (2019).
8. Melamed, Z. et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat. Neurosci.* **22**, 180–190 (2019).
9. Prudencio, M. et al. Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia. *J. Clin. Invest.* **130**, 6080–6092 (2020).
10. Brown, A.-L. et al. TDP-43 loss and ALS-risk SNPs drive missplicing and depletion of *UNC13A*. *Nature* **603**, 131–137 (2022).
11. Ma, X. R. et al. TDP-43 represses cryptic exon inclusion in the FTD-ALS gene *UNC13A*. *Nature* **603**, 124–130 (2022).
12. Seddighi, S. et al. Mis-spliced transcripts generate de novo proteins in TDP-43-related ALS/FTD. *Sci. Transl. Med.* **16**, eadg7162 (2024).
13. Irwin, K. E. et al. A fluid biomarker reveals loss of TDP-43 splicing repression in presymptomatic ALS-FTD. *Nat. Med.* **30**, 382–393 (2024).
14. Gittings, L. M. et al. Cryptic exon detection and transcriptomic changes revealed in single-nuclei RNA sequencing of C9ORF72 patients spanning the ALS-FTD spectrum. *Acta Neuropathol.* **146**, 433–450 (2023).
15. Baughn, M. W. et al. Mechanism of *STMN2* cryptic splice-polyadenylation and its correction for TDP-43 proteinopathies. *Science* **379**, 1140–1149 (2023).
16. Akiyama, T., Koike, Y., Petrucelli, L. & Gitler, A. D. Cracking the cryptic code in amyotrophic lateral sclerosis and frontotemporal dementia: towards therapeutic targets and biomarkers. *Clin. Transl. Med.* **12**, e818 (2022).
17. Mitschka, S. & Mayr, C. Context-specific regulation and function of mRNA alternative polyadenylation. *Nat. Rev. Mol. Cell Biol.* **23**, 779–796 (2022).
18. Passmore, L. A. & Collier, J. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat. Rev. Mol. Cell Biol.* **23**, 93–106 (2022).

19. Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **18**, 18–30 (2017).
20. Rot, G. et al. High-resolution RNA maps suggest common principles of splicing and polyadenylation regulation by TDP-43. *Cell Rep.* **19**, 1056–1067 (2017).
21. Imaizumi, K., Ideno, H., Sato, T., Morimoto, S. & Okano, H. Pathogenic mutation of TDP-43 impairs RNA processing in a cell type-specific manner: implications for the pathogenesis of ALS/FTLD. *eNeuro* **9**, ENEURO.0061-22.2022 (2022).
22. Modic, M. et al. Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. *Mol. Cell* **74**, 951–965 (2019).
23. Ayala, Y. M. et al. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).
24. Prudencio, M. et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175–1182 (2015).
25. McKeever, P. M. et al. Single-nucleus multiomic atlas of frontal cortex in amyotrophic lateral sclerosis with a deep learning-based decoding of alternative polyadenylation mechanisms. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.12.22.573083> (2023).
26. Liu, E. Y. et al. Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. *Cell Rep.* **27**, 1409–1421 (2019).
27. Wang, R. & Tian, B. APALyzer: a bioinformatics package for analysis of alternative polyadenylation isoforms. *Bioinformatics* **36**, 3907–3909 (2020).
28. Ha, K. C. H., Blencowe, B. J. & Morris, Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.* **19**, 45 (2018).
29. Dafinca, R. et al. Impairment of mitochondrial calcium buffering links mutations in C9ORF72 and TARDBP in iPS-derived motor neurons from patients with ALS/FTD. *Stem Cell Rep.* **14**, 892–908 (2020).
30. Herrmann, C. J. et al. PolyASite 2.0: a consolidated atlas of polyadenylation sites from 3' end sequencing. *Nucleic Acids Res.* **48**, D174–D179 (2020).
31. Wang, R., Nambiar, R., Zheng, D. & Tian, B. PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res.* **46**, D315–D319 (2018).
32. Ferrari, R. et al. A genome-wide screening and SNPs-to-genes approach to identify novel genetic risk factors associated with frontotemporal dementia. *Neurobiol. Aging* **36**, 2904 (2015).
33. Bryce-Smith, S. et al. TDP-43 loss induces extensive cryptic polyadenylation in ALS/FTD. *Nat. Neurosci.* <https://doi.org/10.1038/s41593-025-02050-w> (2025).
34. Zhao, W. et al. POSTAR3: an updated platform for exploring post-transcriptional regulation coordinated by RNA-binding proteins. *Nucleic Acids Res.* **50**, D287–D294 (2022).
35. Linder, J., Koplik, S. E., Kundaje, A. & Seelig, G. Deciphering the impact of genetic variation on human polyadenylation using APARENT2. *Genome Biol.* **23**, 232 (2022).
36. Arnold, F. J. et al. TDP-43 dysregulation of polyadenylation site selection is a defining feature of RNA misprocessing in amyotrophic lateral sclerosis and frontotemporal dementia. *J. Clin. Invest.* **135**, e182088 (2025).
37. Figley, M. D., Bieri, G., Kolaitis, R.-M., Taylor, J. P. & Gitler, A. D. Profilin 1 associates with stress granules and ALS-linked mutations alter stress granule dynamics. *J. Neurosci.* **34**, 8083–8097 (2014).
38. Bento-Abreu, A. et al. Elongator subunit 3 (ELP3) modifies ALS through tRNA modification. *Hum. Mol. Genet.* **27**, 1276–1289 (2018).
39. Simpson, C. L. et al. Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration. *Hum. Mol. Genet.* **18**, 472–481 (2009).
40. Volkening, K., Leystra-Lantz, C., Yang, W., Jaffee, H. & Strong, M. J. Tar DNA binding protein of 43 kDa (TDP-43), 14-3-3 proteins and copper/zinc superoxide dismutase (SOD1) interact to modulate NFL mRNA stability. Implications for altered RNA processing in amyotrophic lateral sclerosis (ALS). *Brain Res.* **1305**, 168–182 (2009).
41. Gaetani, L. et al. Neurofilament light chain as a biomarker in neurological disorders. *J. Neurol. Neurosurg. Psychiatry* **90**, 870–881 (2019).
42. Rojas, J. C. et al. Plasma neurofilament light for prediction of disease progression in familial frontotemporal lobar degeneration. *Neurology* **96**, e2296–e2312 (2021).
43. Strong, M. J. et al. TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. *Mol. Cell. Neurosci.* **35**, 320–327 (2007).
44. Ishtiaq, M., Campos-Melo, D., Volkening, K. & Strong, M. J. Analysis of novel NEFL mRNA targeting microRNAs in amyotrophic lateral sclerosis. *PLoS ONE* **9**, e85653 (2014).
45. Lim, Y. W., James, D., Huang, J. & Lee, M. The emerging role of the RNA-binding protein SFPQ in neuronal function and neurodegeneration. *Int. J. Mol. Sci.* **21**, 7151 (2020).
46. Gordon, P. M., Hamid, F., Makeyev, E. V. & Houart, C. A conserved role for the ALS-linked splicing factor SFPQ in repression of pathogenic cryptic last exons. *Nat. Commun.* **12**, 1918 (2021).
47. Luisier, R. et al. Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. *Nat. Commun.* **9**, 2010 (2018).
48. Ishigaki, S. et al. Altered tau isoform ratio caused by loss of FUS and SFPQ function leads to FTLD-like phenotypes. *Cell Rep.* **18**, 1118–1131 (2017).
49. Ishigaki, S. et al. Aberrant interaction between FUS and SFPQ in neurons in a wide range of FTLD spectrum diseases. *Brain* **143**, 2398–2405 (2020).
50. Van Deerlin, V. M. et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat. Genet.* **42**, 234–239 (2010).
51. Chang, A. et al. Homotypic fibrillization of TMEM106B across diverse neurodegenerative diseases. *Cell* **185**, 1346–1355 (2022).
52. Jiang, Y. X. et al. Amyloid fibrils in disease FTLD-TDP are composed of TMEM106B not TDP-43. *Nature* **605**, 304–309 (2022).
53. Schweighauser, M. et al. Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature* **605**, 310–314 (2022).
54. Fan, Y. et al. Generic amyloid fibrillation of TMEM106B in patient with Parkinson's disease dementia and normal elders. *Cell Res.* **32**, 585–588 (2022).
55. Chen-Plotkin, A. S. et al. *TMEM106B*, the risk gene for frontotemporal dementia, is regulated by the microRNA-132/212 cluster and affects progranulin pathways. *J. Neurosci.* **32**, 11213–11227 (2012).
56. Malnar Črnigoj, M. et al. Phenylalanine-tRNA aminoacylation is compromised by ALS/FTD-associated C9orf72 C4G2 repeat RNA. *Nat. Commun.* **14**, 5764 (2023).
57. Salazar, A. N. et al. An AluYb8 mobile element characterises the risk haplotype of the *TMEM106B* locus associated with neurodegeneration. *Alzheimers Dement.* **20**, e090857 (2024).
58. Rodney, A., Karanjeet, K., Benzow, K. & Koob, M. D. A common Alu insertion in the 3'UTR of *TMEM106B* is associated with risk of dementia. *Alzheimers Dement.* **20**, 5071–5077 (2024).
59. Chemparathy, A. et al. A 3'UTR insertion is a candidate causal variant at the *TMEM106B* locus associated with increased risk for FTLD-TDP. *Neurol. Genet.* **10**, e200124 (2024).

60. Marks, J. D. et al. TMEM106B core deposition associates with TDP-43 pathology and is increased in risk SNP carriers for frontotemporal dementia. *Sci. Transl. Med.* **16**, eadf9735 (2024).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the

Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Methods

Cell culture

HEK293T cells were purchased from American Type Culture Collection (CRL-3216) and authenticated by American Type Culture Collection. HEK293 cells were maintained in Dulbecco's Modified Eagle Medium (1×) + GlutaMax-1 media (Gibco, 10564011) with 10% FBS (Gibco, 16000044) and 100 U ml⁻¹ penicillin–streptomycin (Gibco, 15140122). Cells were tested to be negative for mycoplasma using VENORGEN mycoplasma detection kit (Sigma-Aldrich, MP0025).

Stem cell maintenance and differentiation into iNeurons

Human embryonic stem cells (hESCs; H1) were purchased from WiCell (WA01), authenticated by WiCell. H1 cells were maintained in mTeSR1 Plus media (STEMCELL Technologies, 100-0276) with 50 U ml⁻¹ penicillin–streptomycin (Gibco, 15140122) on plates coated with Matrigel (Corning, 354230). hESCs were fed every 2 days and split every 4–7 days using ReLeSR (STEMCELL Technologies, 100-0483) according to the manufacturer's instructions. The differentiation of hESCs to neurons by forcing *NGN2* overexpression was carried out as previously described⁶¹. In brief, cells were transduced with a Tet-On induction system to drive the expression of the transcription factor *NGN2*. Cells were dissociated on day 3 of differentiation and replated on poly-D-lysine and laminin double-coated tissue culture plates in Neurobasal medium (Thermo Fisher Scientific, 21103049) containing neurotrophic factors, brain-derived neurotrophic factor and glial cell line-derived neurotrophic factor (R&D Systems).

TDP-43 KD in iNeurons

After 7 days of being cultured for differentiation, iNeurons were transduced with lentivirus expressing scrambled shRNA or TDP-43 shRNA as previously described⁶¹, cultured for additional 7 days or 12 days, and then collected for downstream analyses. The KD efficiency was assessed by western blotting.

Immunoblotting

After 7 days of TDP-43 KD, cells were lysed at 4 °C for 15 min in ice-cold RIPA buffer (Sigma-Aldrich, R0278) supplemented with a protease inhibitor cocktail (Thermo Fisher Scientific, 78429) and a phosphatase inhibitor cocktail (Thermo Fisher Scientific, 78426). After pelleting lysates at 20,000g for 15 min at 4 °C, the supernatant was used for bicinchoninic acid (Invitrogen, 23225) assays to determine protein concentrations. Unless stated otherwise, 5–10 µg of protein lysates from each sample were denatured for 10 min at 70 °C in LDS sample buffer (Invitrogen, NP0008) containing 2.5% 2-mercaptoethanol (Sigma-Aldrich). These samples were loaded onto 4–12% Bis–Tris Mini gels (Thermo Fisher Scientific, NP0335BOX) for gel electrophoresis and then transferred onto 0.45-µm nitrocellulose membranes (Bio-Rad, 162-0115) using the semi-dry transfer method (Bio-Rad Trans-Blot Turbo Transfer System, 1704150) or at 100 V for 2 h at 4 °C using the wet transfer method (Bio-Rad Mini Trans-Blot Electrophoretic Cell, 170-3930). Membranes were blocked in EveryBlot Blocking Buffer (Bio-Rad, 12010020) or 5% nonfat dry milk in Tris-buffered saline with Tween 20 for 1 h then incubated overnight at 4 °C in blocking buffer containing antibodies against TMEM106B (1:500; Cell Signaling Technology, 93334), TDP-43 (1:1,500; Proteintech, 10782-2-AP), ELP1 (1:500; Cell Signaling Technology, 5071S), ELP3 (1:1,000; Proteintech, 17016-1-AP), NEFL (1:1,000; Thermo Fisher Scientific, MA5-14981), SFPQ (EPRI1847) (1:1,000; Abcam, ab177149), GAPDH (1:2,000; Sigma-Aldrich, G8795) or GAPDH (D16H11) XP (1:1,000; Cell Signaling Technology, 8884S). Membranes were subsequently incubated in blocking buffer containing horseradish peroxidase-conjugated antimouse IgG (H + L; 1:5,000; Thermo Fisher Scientific, 62-6520) or horseradish peroxidase-conjugated antirabbit IgG (H + L; 1:5,000; Life Technologies, 31462) for 1 h. The Amersham ECL Prime kit (Cytiva, RPN2232) or SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific,

34094) was used to develop blots, which were then imaged using ChemiDox XRS+ System (Bio-Rad). The intensity of bands was quantified using Fiji, and then normalized to the corresponding controls.

Immunofluorescence

After 7 days of being cultured for differentiation, cells were fixed with 4% formaldehyde (Electron Microscopy Sciences, 15710) in 1× PBS at room temperature for 15 min, washed four times with 1× PBS and permeabilized with 0.2% Triton X-100 in the blocking buffer (5% normal goat serum in 1× PBS) at room temperature for 1 h. After permeabilization and blocking, cells were stained in blocking buffer containing antibodies against VAT1L (1:500; Thermo Fisher Scientific, PA5-98934), NeuN (1:500; EMD Millipore, MAB377) and Tuj1 (1:1,000; Covance, MMS-435P) in a cold room overnight. After primary antibody staining, cells were washed four times with 1× PBS, twice with the blocking buffer and incubated with secondary antibodies in the blocking buffer containing Hoechst 33342 (1:1,000) at room temperature for 1 h. After the 1 h incubation, cells were washed four times with 1× PBS, and imaged on a Leica DMI6000 B with a ×20 objective and an HAMAMATSU ORCA-flash 4.0 camera.

Immunoblotting of TMEM106B

After 12 days of TDP-43 KD, cells were lysed and normalized as described above. To detect the impact of TDP-43 KD on the dimer level of TMEM106B, normalized cell lysates were mixed with 4× Laemmli buffer (Bio-Rad) containing 5% 2-mercaptoethanol (Sigma-Aldrich) and loaded directly onto 4–20% Tris–Glycine mini gels (Thermo Fisher Scientific, XP04205BOX) for gel electrophoresis on ice. To detect the temperature sensitivity of TMEM106B dimer, cell lysates were incubated at 4 °C, 37 °C, 70 °C and 85 °C for 10 min and then loaded onto a 4–20% Tris–Glycine mini gel for gel electrophoresis on ice. After electrophoresis, samples were transferred onto 0.45-µm polyvinylidene difluoride membranes (Bio-Rad, 162-0115) at 250 mA for 2 h using the wet transfer method (Bio-Rad Mini Trans-Blot Electrophoretic Cell, 170-3930). Membranes were blocked in 5% nonfat dry milk in Tris-buffered saline with Tween 20 for 1 h and then incubated overnight at 4 °C in blocking buffer containing the antibody against TMEM106B (1:500; Cell Signaling Technology, 93334). The secondary antibody, imaging and quantitation are the same as described above.

Total RNA extraction from iNeurons

Total RNA was extracted using TRIzol according to the manufacturer's instructions. Total RNA was then treated with Turbo DNase (Thermo Fisher Scientific, AM2238) and cleaned using Zymo's clean and concentration columns. The quality of RNA was examined on a High Sensitivity RNA ScreenTape (Agilent Technologies, TapeStation).

qRT-PCR from iNeurons

Total RNA (500 ng) was reverse transcribed to cDNA using the PrimeScript RT reagent Kit with gDNA Eraser (Takara, RRO47A). qPCR was carried out using PowerUp SYBR Green Master Mix kit and detected by the QuantStudio3 system (Thermo Fisher Scientific). Primers are listed in Supplementary Table 7.

Competition PCR from iNeurons

Total RNA (250 ng) was reverse transcribed to cDNA using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, RRO47A). Competition PCR was carried out using the GoTaq DNA polymerase (Promega, M3001) using one forward primer and two reverse primers—one targeting the canonical junction of intron 9 and the other targeting the alternative junction of intron 9. The resulting PCR products were visualized on a 2% agarose gel. Primers are listed in Supplementary Table 7.

Total RNA-seq

Total RNA from scrambled shRNA or TDP-43 shRNA-treated cells was used to construct RNA-seq libraries using the SMARTer Stranded Total

RNA-Seq Kit v2–Pico Input Mammalian kit (Takara, 634411), according to the manufacturer's instructions. The resulting libraries were quantified, pooled and sequenced on a NextSeq 500 machine using the 150-cycle high-output kit in a 75 bp paired-end mode (Illumina).

Gene expression analysis

Adapters in FASTQ files were trimmed using fastp. The adapter-trimmed FASTQ files were used for differential gene expression analysis using Salmon and DESeq2.

Splicing analysis

The adapter-trimmed FASTQ files were mapped to the human genome (hg38) following ENCODE's recommended settings using STAR. The uniquely mapped, properly paired reads were then used for splicing analysis using LeafCutter. Cryptic splicing events were called by LeafCutter.

APA analysis using RNA-seq data

Both in-house RNA-seq and publicly available datasets were used for the analysis. The publicly available datasets were downloaded from GEO or SRA (GSE126542, GSE121569, GSE147544, GSE196144 and ERP126666). The adapter-trimmed FASTQ files were mapped to the human genome (hg38) following ENCODE's recommended settings using STAR with GENCODE v42 transcriptome. The unique-mapping, properly paired reads were then used for APA detection by either APALyzer or QAPA and then for differential analysis by DEXSeq. APALyzer uses polyA DB3 (https://exon.apps.wistar.org/PolyA_DB/) and QAPA uses PolyASite 2.0 (<https://polyasite.unibas.ch/>) as polyA site references for detecting APA events. The Δ PUI used in APALyzer is defined as the difference in the \log_2 -transformed distal polyA site usage between two conditions.

3' end-seq

To comprehensively map polyadenylation and quantify APA, three control samples and three TDP-43 KD samples were used for 3' end-seq. For each sample, 500 ng of total RNA was used to construct the 3' end-seq library using the Quantseq REV kit from Lexogen, according to the manufacturer's instructions. The resulting 3' end-seq libraries were quantified by Qubit, checked for library sizes on a D1000 high-sensitivity chip (Agilent Technologies, TapeStation), pooled and sequenced on a NextSeq 500 or NextSeq 2000 machine using the 150-cycle high-output kit in a 75 bp paired-end mode (Illumina).

APA analysis using 3' end-seq data

The sequenced 3' end-seq libraries were adapter trimmed and quality filtered using BBduk. The filtered reads were mapped to the human genome (hg38) using STAR with GENCODE v42 transcriptome, extracted for unique-mapping, properly paired reads using Samtools, and deduplicated using Picard tools. To remove mis-primed reads, we discarded reads that were mapped upstream of six consecutive As, upstream of a 10-bp region with at least 50% of As or upstream of one of the six pentamers—AAAAA, GAAAA, AGAAA, AAGAA, AAAGA and AAAAG—that might form the G–T wobble pair during cDNA synthesis. The resulting filtered reads were analyzed using a modified version of long-read alternative polyadenylation analysis (LAPA)⁶² that used the default setting, except that it required a replication rate cutoff at 0.75 and could identify polyA site-containing reads mapped to the reverse strand. We further filtered for identified polyA sites if they were present in all samples in at least one condition, had at least ten average reads across samples, had at least 5% usage rate in at least one condition and were not located in intergenic regions. This stringent filtering may remove false negative reads (that is, reads misclassified as mis-primed reads); however, this filtering allows us to capture most reproducible and substantial APA changes caused by TDP-43 KD. After this filtering, genes with at least two polyA sites were used for APA analysis. We used StringTie to assemble transcript isoforms from matched RNA-seq data

and then used the assembled isoforms to help assign polyA sites that are located downstream of annotated transcript 3' ends. The change in APA is defined as the usage difference in the distal *p*-aminosalicylic acid usage rate (Δ dPUR) between TDP-43 KD and control samples. The APA change was considered substantial if $|\Delta$ dPUR| > 0.05 with an adjusted *P* value of <0.05. If genes had more than two polyA sites, they were filtered for two polyA sites with the two smallest adjusted *P* values and used for plotting and downstream sequence feature analysis (Supplementary Table 4). PolyA sites were considered cryptic if their usage was <10% under the control condition but >10% under the TDP-43 KD condition, and their usage increase was \geq 10% (Supplementary Table 3). Cryptic splicing events detected in RNA-seq data and 3' end-seq data were used to search for coordinated cryptic polyadenylation events.

Alignment of read ends across transcripts

Using DeepTools, the ends of filtered reads were calculated for normalized coverage (RPKM) across protein-coding genes and lncRNA genes, processed with computeMatrix, and then plotted using plotHeatmap.

Analysis of polyA signals

The 60-nt region upstream of each polyA identified by LAPA was examined for any known polyA signals previously documented³⁰. The frequency of each polyA signal was then plotted for both annotated and new polyA sites. Polyadenylation signals that were not AATAA or ATAAA were categorized as the APA signal. Subsequently, the distances from these polyA signals to the corresponding polyA sites were plotted for both annotated and new polyA sites.

The position and strength of TDP-43-binding site

The 1,000-nt region surrounding each TDP-43-regulated APA site was analyzed for overlaps with TDP-43-binding sites identified through CLIP-seq. We calculated the distance from each overlapping TDP-43-binding site to its corresponding APA site. Only those TDP-43-binding sites with the shortest distances were retained for further analysis. For these selected binding sites, we estimated their strengths based on the GU content within a 50-nt region surrounding each site.

Hexamer analysis

To identify enriched and depleted hexamers in specific regions surrounding TDP-43-regulated APA sites, *k*-mer analysis was performed using the transite package in R.

Calculation of polyA site score

For each polyA site identified in the 3' end-seq libraries, a 205 bp sequence centered at the site was extracted and used to calculate the polyA site score using Aparent2, which was then converted to the log odds ratio.

Inhibition of the NMD

iNeurons were treated with different concentrations of 11j (MedChem Express, HY-124719), an SMG-1 inhibitor or an equal volume of dimethyl sulfoxide for 24 h and collected for western blot to confirm that phosphorylation of UPF1 was inhibited. Western blot was performed as described above with wet transfer and proteins were detected using the following antibodies: UPF1 (1:1,000; Abcam, ab109363), pUPF1 (1:1,000; EMD Millipore, 07-1016) and GAPDH (1:2,000; Sigma-Aldrich, G8795). Samples treated with 0.01 μ m 11j were further used for qRT-PCR with primers listed in Supplementary Table 7.

Pulse-chase experiment using actinomycin-D

After 12 days of TDP-43 KD, iNeurons were treated with 20 μ m of actinomycin-D (Sigma-Aldrich, A1410) and collected at time points of 0, 1, 3 and 6 h post-treatment for total RNA extraction using TRIzol (Thermo Fisher Scientific, 15596026). The resulting RNA was measured by Qubit (Thermo Fisher Scientific), normalized to the same concentration, and

used for qRT-PCR with primers listed in Supplementary Table 7. The normalized RNA levels over time were fitted with an exponential decay model using nls function in R ($y - y_0 \times \exp(-k \times \text{time})$) with starting values of y_0 at 1 and k at 0.01) to calculate the decay rate (r), which were then used to estimate the half life with $\ln(2)/r$.

Evaluation of *SFPQ* 3' UTR length, *SFPQ* cryptic splicing and *TMEM106B* 3' UTR length in FTLD-TDP brain samples by qRT-PCR

Postmortem brain tissues from patients with FTLD-TDP and cognitively normal control individuals were obtained from the Mayo Clinic Florida Brain Bank. Diagnosis was independently ascertained by trained neurologists and neuropathologists upon neurological and pathological examinations, respectively. Written informed consent was given by all participants or authorized family members, and all protocols were approved by the Mayo Clinic Institutional Review Board and Ethics Committee. Our study cohort included a total of 83 postmortem cases classified into the following two main groups: healthy controls ($n = 53$) and FTD cases ($n = 221$). A summary of patient data is included in Supplementary Table 8. RNA was extracted from the frontal cortices of healthy and FTD patients following the manufacturer's protocol using the RNeasy Plus Mini Kit (Qiagen) and as previously described^{9,11,63}. Up to three cuts of the sample were used for extraction and only the high-quality RNA samples were processed for downstream analysis. RNA concentration was measured by using Nanodrop technologies (Thermo Fisher Scientific) and the RNA integrity number was evaluated by the Agilent 2100 Bioanalyzer (Agilent Technologies, TapeStation). Subsequently, 500 ng of the total RNA extracted was reverse transcribed into cDNA using the High-Capacity cDNA Transcription Kit (Applied Biosystems) as per the manufacturer's instructions. cDNA samples, in triplicate, with SYBR GreenER qPCR SuperMix (Invitrogen), were further evaluated for qRT-PCR on a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). Relative quantification of levels of the long *SFPQ* 3' UTR, the total *SFPQ* 3' UTR, the long *TMEM106B* 3' UTR and total *TMEM106B* was determined using the $\Delta\Delta C_t$ method and normalized to two endogenous controls, *GAPDH* and *RPLPO*. All statistical analyses were performed using the GraphPad Prism 10 (GraphPad software). For comparison of the frontal cortex RNA levels between the healthy and FTD cases, the Mann-Whitney test was used. Primers are listed in Supplementary Table 8.

Construction of luciferase reporters

The short *TMEM106B* 3' UTR and the long *TMEM106B* 3' UTR were amplified from human genomic DNA (HI) and then cloned into the pmirGLO Dual-Luciferase Vector (Promega, E1330) using Gibson assembly. The proximal polyA site in the long *TMEM106B* 3' UTR was then mutated using Gibson assembly. Mutations that disrupt the proximal polyA site were identified using Aparent2. The polyA signal was identified using Aparent2. All of the reporters were confirmed to have correct sequences by whole-plasmid sequencing, and their sequences can be found in Supplementary Table 9.

Overexpression of exogenous *TMEM106B*

The lentiviral plasmid containing FLAG-tagged *TMEM106B* lacking its 3' UTR and the corresponding lentivirus were made by VectorBuilder. Their sequences can be found in Supplementary Table 9. After 7 days of TDP-43 KD, iNeurons were transduced with lentivirus containing either the FLAG-tagged *TMEM106B* construct or the EGFP control and collected 5 days post-transduction for dimer-level analysis using western blot. The western blot was carried out as described in the immunoblotting of *TMEM106B*.

Measurement of luciferase activity

For *TMEM106B* reporters, HEK293T cells were plated on a 96-well plate or a 24-well plate and transfected with dual-luciferase reporters

that have no insert, the short *TMEM106B* 3' UTR or the long *TMEM106B* 3' UTR using Lipofectamine 3000 (Thermo Fisher Scientific, L3000001). Two days later, the transfected cells on a 96-well plate were measured for the luciferase activities using the Dual-Glo Luciferase Assay System (Promega, E2920), according to the manufacturer's instructions; the transfected cells on a 24-well plate were used for qRT-PCR as described above.

Quantitation and statistical analysis

All quantification and statistical analyses were done in R. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications^{11,20}. Experiments were generally not anonymized or randomized during analysis. Data distributions in Fig. 3a,c-e and Supplementary Fig. 2o were tested for normality and other data distributions were not formally tested, but all data points were plotted for every experiment. Analysis details can be found in figure legends, Methods and 'Main'. Genome tracks were prepared using IGV. All plots were prepared using tidypplots, ggplot2, ggpubr, patchwork and ggrepel in R.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequencing data generated in this study are available on GEO (GSE252892). The following publicly available GEO data are used in this study: GSE126542, GSE121569, GSE147544, GSE196144 and ERP126666. The human genome (hg38) and transcriptome references (v42) are available on GENCODE (https://www.genencodegenes.org/human/release_42.html). Source data are included with this paper and available on Zenodo (<https://doi.org/10.5281/zenodo.15637268>)⁶⁴. Source data are provided with this paper.

Code availability

Codes to generate each figure are available on Zenodo (<https://doi.org/10.5281/zenodo.15637268>)⁶⁴ and GitHub (https://github.com/koryant/tdp43_apa).

References

- Bieri, G. et al. LRRK2 modifies α -syn pathology and spread in mouse models and human neurons. *Acta Neuropathol.* **137**, 961–980 (2019).
- Çelik, M. H. & Mortazavi, A. Analysis of alternative polyadenylation from long-read or short-read RNA-seq with LAPA. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.11.08.515683> (2022).
- Estades Ayuso, V. et al. TDP-43-regulated cryptic RNAs accumulate in Alzheimer's disease brains. *Mol. Neurodegener.* **18**, 57 (2023).
- Zeng, Y. & Gitler, A. TDP-43 nuclear loss in FTD/ALS causes widespread alternative polyadenylation changes (version V2). *Zenodo* <https://doi.org/10.5281/zenodo.15637268> (2025).

Acknowledgements

We thank the members of the Gitler lab and the Petrucelli lab for their helpful discussions and comments on the paper. We thank Z. Chen for helping with Aparent2. We also acknowledge the funding support from the following sources: a postdoctoral scholar award from the Phil and Penny Knight Initiative for Brain Resilience, Stanford University (to Y.Z.); a grant from the Larry L. Hillblom Foundation (to Y.Z.); the National Institutes of Health (NIH) Training Program in Aging Research (2T32AG047126-06A1 to T.A.); a grant from the Takeda Science Foundation (to T.A.); FightMND (to S.L.R.); a Stanford School of Medicine Dean's postdoctoral fellowship (to C.L.);

the Milton Safenowitz Postdoctoral Fellowship Program (to C.G.); the NIH (U54NS123743, R35NS097273 and P01NS084974) and Target ALS (to L.P.); the NIH (U54NS123743 and RF1NS120992 to M.P.); and the NIH (R35NS137159, U54NS123743 and R01AG064690) and Target ALS (to A.D.G.). A.D.G. is an investigator at Chan Zuckerberg Biohub—San Francisco. Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. Part of the graphics was created with [BioRender.com](https://www.biorender.com) (Fig. 6e and Supplementary Fig. 1c). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Y.Z. and A.D.G. conceptualized the study, wrote the original draft, and revised and edited the paper with input from all authors. Y.Z., A.L., T.A., S.L.R. and V.M.J. designed the methodology and conducted the formal analysis, validation and visualization. Y.Z., A.L., T.A., S.L.R., V.M.J., C.L., O.S., C.G. and A.C. performed data curation and investigation. Y.Z. developed the analysis pipeline. Y.Z., L.P. and A.D.G. managed

project administration. M.P., D.W.D., L.P. and A.D.G. provided resources, supervision and funding acquisition.

Competing interests

A.D.G. is a scientific founder of Maze Therapeutics, Trace Neuroscience and Lyterian Therapeutics. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-025-02049-3>.

Correspondence and requests for materials should be addressed to Yi Zeng or Aaron D. Gitler.

Peer review information *Nature Neuroscience* thanks Junjie Guo, Jemeen Sreedharan, Bin Tian and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|---|
| Data collection | No software or code used in the the data collection stage. |
| Data analysis | fastp/0.22.0, Salmon/1.10.0, DESeq2/1.38.3, BBduk/39.01, STAR/2.7.10a, regtools/1.0.0, LeafCutter/v0.2.1, APALyzer/1.12.0, QAPA/1.3.3, LAPA/0.0.4, DEXSeq/1.44.0, Samtools/1.17, Picard/2.27.4, StringTie/2.2.1, Deeptools/3.5.1, Python/3.10.11, R/4.2.0, GraphPad Prism 10, Aparent2, snakemake/7.32.4. The code are available on Zenodo (https://doi.org/10.5281/zenodo.14214058) and on Github (https://github.com/koryant/tdp43_ap). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The sequencing data generated in this study is available at GEO: GSE252892. The following publicly available GEO data are used in this study: GSE126542,

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Biological sex of the de-identified patient samples are included as summary in supplementary table 8.
Reporting on race, ethnicity, or other socially relevant groupings	Relevant information on race and ethnicity are included in supplementary table 8.
Population characteristics	We have everything listed in supplementary table 8.
Recruitment	Written informed consent was obtained before study entry from all subjects or their legal kin if they were unable to give written consent and biological samples were obtained with Mayo Clinic Institutional Review Board (IRB) approval.
Ethics oversight	Mayo Clinic Institution Review Board and Ethics committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed; for in vitro experiments, sample sizes were chosen such that statistical significance could be confidently established and are similar to previous studies on similar topics (Ma et al. 2022, PMID:35197626; Rot et al., 2017, PMID: 28467899); or analysis using patient data, sample sizes were chosen based on data availability.
Data exclusions	No data from in vitro experiments were excluded. Minimal patient data were excluded based on criteria described in Methods.
Replication	All attempts at replication were successful. At least three independent replications are performed for each experiment except for Figure s3r, which had 2 replicates. Key findings are also validated by orthogonal experiments.
Randomization	Experimental conditions were not randomized as they were predetermined to access the impact of specific treatments on alternative polyadenylation, splicing, RNA levels, or protein levels.
Blinding	The researchers were not blinded to the identity of the samples but findings have been replicated by independent researchers within the lab and across multiple laboratories. Key findings are also validated by orthogonal experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants |

Methods

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used	<p>TMEM106B (clone EZH7Z, 1:500, Cell Signaling Technology 93334), TDP-43 (1:1,500, Proteintech 10782-2-AP), ELP1 (1:500, Cell Signaling Technology 5071S), ELP3 (1:1000, Proteintech 17016-1-AP), NEFL (clone T.400.5, 1:1000, Thermo Fisher MA5-14981), GAPDH (clone 71.1, 1:2,000, Sigma-Aldrich G8795), SFPQ (clone EPR11847, 1:1000; Abcam, ab177149) VAT1L (1:500; Thermo Scientific, PA5-98934) NeuN (clone A60, 1:500; EMD Millipore, MAB377) Tuj1 (1:1000; Covance, MMS-435P) UPF1 (clone EPR4681, 1:1,000; Abcam ab109363) pUPF1 (1:1,000; EMD Millipore, 07-1016)</p>
Validation	<p>All commercial antibodies were validated by their corresponding manufacturers, as follows:</p> <ul style="list-style-type: none"> - TMEM106B: https://www.cellsignal.com/products/primary-antibodies/tmem106b-e7h7z-rabbit-mab/93334. The specificity of this antibody was also validated independently in house by knocking down TMEM106B in iNeurons, HEK293 cells or using a TMEM106B KO cell line. - TDP-43: https://www.ptglab.com/products/TARDBP-Antibody-10782-2-AP.htm. The specificity of this antibody was validated independently in house by knocking down TDP-43 in various cell lines or using a TDP-43 KO cell line. - ELP1: https://www.cellsignal.com/products/primary-antibodies/elp1-ikbkap-antibody/5071. The specificity of this antibody was tested by the vendor across multiple cell lines. - ELP3: https://www.ptglab.com/products/ELP3-Antibody-17016-1-AP.htm. The band with expected molecular weight for ELP3 was detected by WB in house. - NEFL: https://www.thermofisher.com/antibody/product/NEFL-Antibody-clone-T-400-5-Monoclonal/MA5-14981. The specificity of this antibody was validated by the vendor in a NEFL KO cell line. - GAPDH: https://www.sigmaaldrich.com/US/en/product/sigma/g8795. The specificity of this antibody was tested by the vendor across multiple cell lines. - SFPQ: https://www.abcam.com/en-us/products/primary-antibodies/sfpq-antibody-epr11847-ab177149. The specificity of this antibody was tested by the vendor across multiple cell lines. - VAT1L: https://www.thermofisher.com/antibody/product/VAT1L-Antibody-Polyclonal/PA5-98934. N/A - NeuN: https://www.sigmaaldrich.com/US/en/product/mm/mab377. The specificity of this antibody was validated by numerous studies, see here: https://www.citeab.com/antibodies/226230-mab377-anti-neun-antibody-clone-a60. - Tuj1: https://www.biolegend.com/ja-jp/products/purified-anti-tubulin-beta-3-tubb3-antibody-11580. The specificity of this antibody was validated by numerous studies. - UPF1: https://www.abcam.com/en-us/products/primary-antibodies/rent1-hupf1-antibody-epr4681-ab109363#. The specificity of this antibody was tested by the vendor across different cell lines and by immunoprecipitation - pUPF1 https://www.sigmaaldrich.com/US/en/product/mm/071016. The specificity of this antibody was tested via immunoprecipitation and phosphorylated peptides.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK293 cells were purchased from purchased from ATCC (CRL-3216). Parental H1 embryonic stem cell line is from WiCell. H1 with inducible NGN2 expression is made in house.
Authentication	HEK293 cells and H1 embryonic stem cell line were authenticated by the vendors using STR analysis.
Mycoplasma contamination	Cells were tested to be negative for mycoplasma using VENORGEM mycoplasma detection kit (Sigma, MP0025).
Commonly misidentified lines (See ICLAC register)	Cell lines were not listed in ICLAC register.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	N/A
Study protocol	N/A
Data collection	The clinical data we used in this study were from tissues donated by deceased patients.
Outcomes	N/A

Plants

Seed stocks

N/A

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.