


RESEARCH ARTICLE

Open Access



# Elevated nuclear TDP-43 induces constitutive exon skipping

Rogger P. Carmen-Orozco<sup>1,2†</sup>, William Tsao<sup>1,2†</sup>, Yingzhi Ye<sup>3†</sup>, Irika R. Sinha<sup>1,2</sup>, Koping Chang<sup>1</sup>, Vickie T. Trinh<sup>1,2</sup>, William Chung<sup>1</sup>, Kyra Bowden<sup>1</sup>, Juan C. Troncoso<sup>1</sup>, Seth Blackshaw<sup>2,4,5</sup>, Lindsey R. Hayes<sup>5</sup>, Shuying Sun<sup>1,2,3</sup>, Philip C. Wong<sup>1,2</sup> and Jonathan P. Ling<sup>1\*</sup> 

## Abstract

**Background** Cytoplasmic inclusions and loss of nuclear TDP-43 are key pathological features found in several neurodegenerative disorders, suggesting both gain- and loss-of-function mechanisms of disease. To study gain-of-function, TDP-43 overexpression has been used to generate in vitro and in vivo model systems.

**Methods** We analyzed RNA-seq datasets from mouse and human neurons overexpressing TDP-43 to explore species specific splicing patterns. We explored the dynamics between TDP-43 levels and exon repression in vitro. Furthermore we analyzed human brain samples and publicly available RNA datasets to explore the relationship between exon repression and disease.

**Results** Our study shows that excessive levels of nuclear TDP-43 protein lead to constitutive exon skipping that is largely species-specific. Furthermore, while aberrant exon skipping is detected in some human brains, it is not correlated with disease, unlike the incorporation of cryptic exons that occurs after loss of TDP-43.

**Conclusions** Our findings emphasize the need for caution in interpreting TDP-43 overexpression data and stress the importance of controlling for exon skipping when generating models of TDP-43 proteinopathy.

<sup>†</sup>Rogger P. Carmen-Orozco, William Tsao and Yingzhi Ye contributed equally to this work.

\*Correspondence:

Jonathan P. Ling  
jling@jhu.edu

<sup>1</sup>Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

<sup>2</sup>Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

<sup>3</sup>Department of Physiology, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

<sup>4</sup>Department of Ophthalmology, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

<sup>5</sup>Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

## Background

Transactivation response element DNA-binding protein 43 (*TARDBP*, TDP-43) is an RNA-binding protein implicated in a variety of age-related disorders such as amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD) [1–3], Alzheimer's disease (AD) [4], limbic-predominant age-related TDP-43 encephalopathy [5], Paget's Disease of Bone [6], and inclusion body myositis [7]. TDP-43 proteinopathy is characterized by large cytoplasmic inclusions and loss of nuclear TDP-43 staining [8], suggesting that both gain- and loss-of-function effects could contribute to disease pathogenesis.

Genetic deletion models have demonstrated that TDP-43 is an essential gene [9–14]. Loss of TDP-43 function has been linked to neurodegeneration via disruptions in the splicing repression of nonconserved cryptic exons [7,



15–29]. Indeed, compelling evidence from recent studies suggests that cryptic exons found in the genes *STMN2* [17–19] and *UNC13A* [22, 23] may contribute significantly to disease pathogenesis. Repression of cryptic exons in these and other genes represents a promising therapeutic strategy for ALS-FTD and other neurodegenerative disorders [29].

By contrast, transgenic models of TDP-43 overexpression also exhibit toxicity, but do not reproduce the cytoplasmic aggregates observed in human disease [30–42]. Despite the absence of cytoplasmic aggregation [43], these TDP-43 transgenic models exhibit a dose-dependent degeneration of cortical and spinal motor neurons [32, 44, 45] that may be linked to excessive levels of nuclear TDP-43. Indeed, TDP-43 autoregulates the stability of its own mRNA by binding to an ultraconserved element in its 3' untranslated region [46–48]. Only a single genomic copy of *Tardbp* is required for survival, as *Tardbp* heterozygotes exhibit the same levels of TDP-43 protein as wildtypes [11]. Collectively, these data indicate that elevated levels of TDP-43 protein are harmful, even in non-disease contexts.

Alternative strategies to model TDP-43 proteinopathy have been developed, involving the use of adeno-associated virus (AAV)-mediated delivery of TDP-43 [49] and the generation of transgenic TDP-43 models that feature mutated nuclear localization signals (NLS), such as rNLS8 [50–53]. The use of AAV delivery bypasses development-specific effects and enables delivery to specific cell types in adult mice, while the animal models that express NLS-deficient TDP-43 enable the modelling of cytoplasmic toxicity [54] that are not recapitulated by wildtype overexpression [37]. These studies indicate a potential unified mechanism of neurodegeneration in which microglia exert a neuroprotective role by phagocytically clearing pathological TDP-43 [49–51], whereas impaired microglial function could worsen neuronal damage and motor impairments. However, it is not known whether the molecular mechanisms that produce the pathological changes in these overexpression models are equivalent to those that occur in human disease. Fully characterizing these mechanisms will be crucial to accurately interpret TDP-43 overexpression models and their relevance to human disease.

Prior work in transgenic mouse models has demonstrated that increased expression of wild-type or mutant TDP-43 in mice caused widespread splicing changes, with differing effects depending on specific exons and TDP-43 variants [55]. While elevated wild-type TDP-43 increased exon skipping for many exons, mutant TDP-43 showed both loss-of-function (reduced exon skipping) and gain-of-function (skipping of new exons) effects in an exon-dependent manner that also depended on levels of transgene expression. Further studies used mouse

mutants carrying point mutations in endogenous *Tardbp* to identify a novel class of exons, termed “skiptic exons”, that are skipped due to gain-of-function effects from these *Tardbp* point mutations [56]. These skiptic exons show high conservation across species and 2 out of 7 skiptic events were validated in fibroblasts from ALS patients with disease-causing *TARDBP* mutations. Although TDP-43 RNA binding was proposed as the likely mediator of splicing gain-of-function, the authors also identified that various point mutations are able to disrupt autoregulation and lead to increased mRNA levels of *Tardbp*, as was reported in a parallel study [57]. These findings highlight the nuanced impact of TDP-43 mutations and expression levels on alternative splicing regulation.

In this work, we demonstrate that excessive levels of nuclear TDP-43 protein, mutant or wildtype, leads to the repression of constitutive exons that are normally incorporated into mRNAs. Sufficiently high levels of TDP-43 overexpression can overwhelm autoregulation, even when nuclear localization-deficient isoforms of TDP-43 [50, 52, 58] are introduced. TDP-43 preferentially binds to long repetitive UG repeats [16, 59], but with higher concentrations in the nucleus, we theorize that TDP-43 can bind to shorter, less optimal UG-containing motifs present at these constitutive exons.

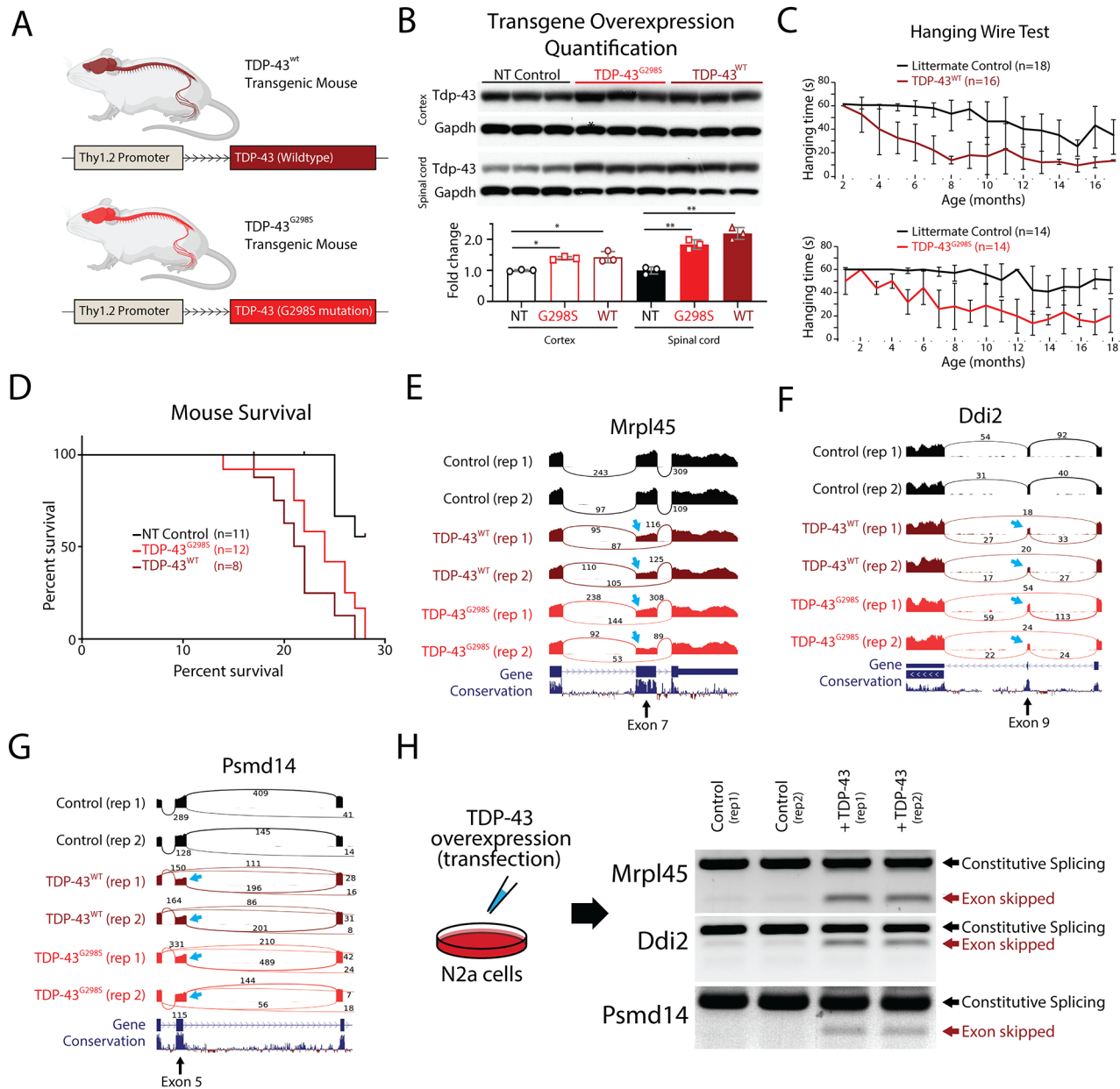
Interestingly, the constitutive exons repressed by excessive TDP-43 are mostly divergent between mouse and human neurons. We also find that, while aberrant constitutive exon skipping can be detected in some human brain samples, constitutive exon skipping does not correlate with disease. By contrast, TDP-43 dependent cryptic exons are found only in human disease tissues or biofluids. Our findings imply that constitutive exon skipping is associated with TDP-43 overexpression, but its direct link to neurodegeneration requires further investigation in vivo. The consequences of skiptic splicing should be considered when developing models of TDP-43 proteinopathy to ensure that these models accurately reflect disease mechanisms.

## Results

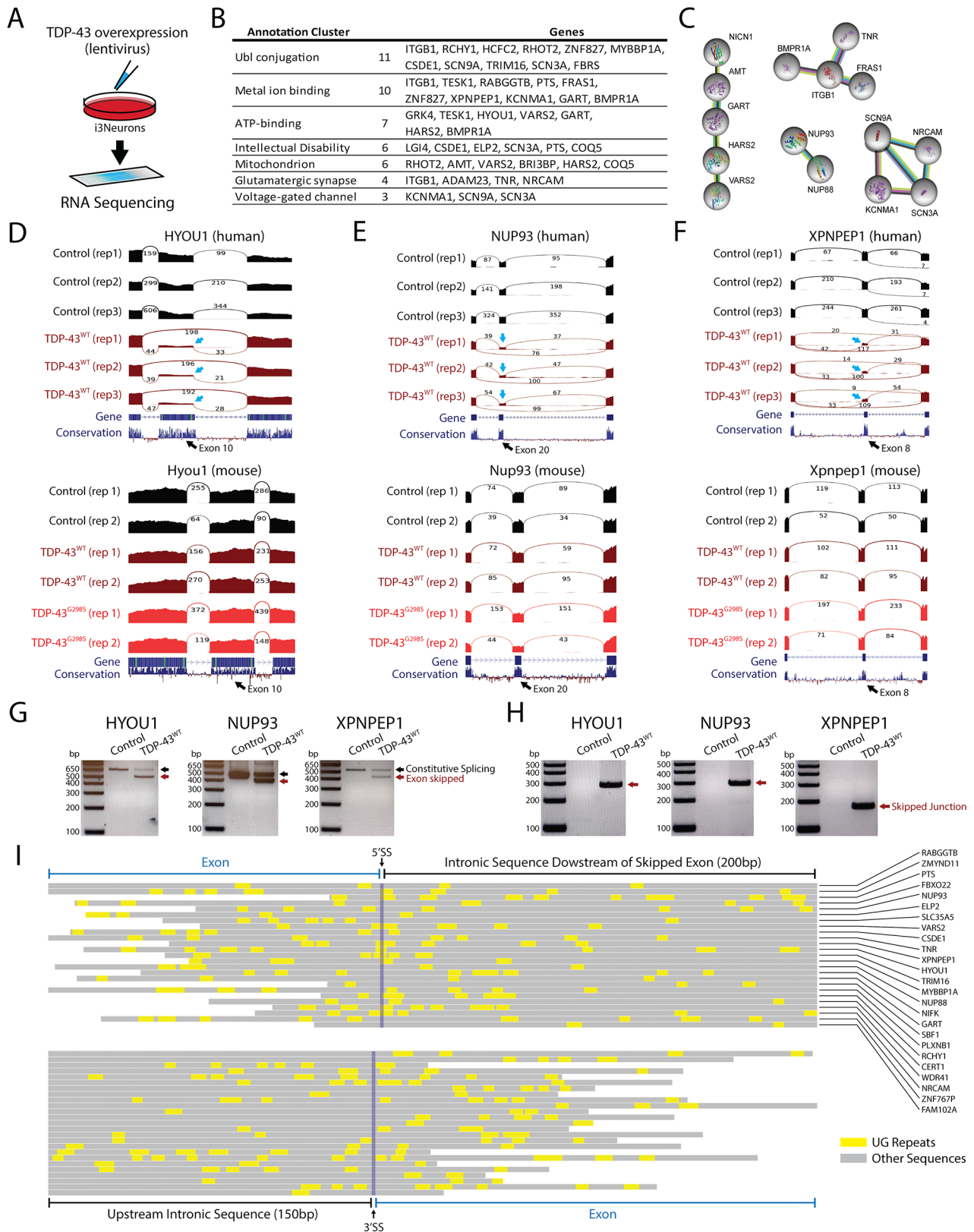
Transgenic models of TDP-43 overexpression exhibit a dose-dependent toxicity [31, 60–62], but do not recapitulate the nuclear clearance and cytoplasmic aggregation that are hallmarks of TDP-43 proteinopathy. TDP-43 overexpression models also do not exhibit any disruptions in cryptic exon repression, as found in models of TDP-43 loss-of-function [15]. These findings suggest that increasing TDP-43 protein to a degree that significantly exceeds physiological levels introduces gain-of-function toxicity that may not be applicable to the pathogenesis of human disease. Therefore, to determine whether slight elevations in TDP-43 levels could produce better models

of disease, we generated transgenic mice that express wildtype human TDP-43 (TDP-43<sup>WT</sup>) or human TDP-43 carrying the G298S mutation (TDP-43<sup>G298S</sup>) that is associated with familial ALS [63], under the control of the Thy1.2 promoter (Fig. 1A). Immunoblotting analysis indicates that total TDP-43 protein levels in these transgenic

models was approximately 50% higher than in non-transgenic controls (Fig. 1B). Indeed, transgenic animals developed a progressive motor deficit that began with hind-limb clench and hemiparesis, eventually leading to end-stage paralysis. Behavioral testing showed marked reduction on hanging time in transgenic lines compared



**Fig. 1** TDP-43 overexpression in mice leads to skipping of constitutive exons. **(A)** Human TDP-43 (TDP-43<sup>WT</sup>) and TDP-43 carrying a G298S mutation (TDP-43<sup>G298S</sup>) were expressed under the weak Thy1.2 promoter in mice. **(B)** TDP-43 levels in the spinal cord and cortex of transgenic mice were compared to control mice using immunoblotting. In both transgenic lines, TDP-43 protein levels were elevated at approximately 1.5x and 1.3x higher in the spinal cord and cortex, respectively (\* $p < 0.05$ , \*\*\* $p < 0.001$ ). **(C)** We measured the hanging time of transgenic mice compared to their littermate controls and found a reduction associated with age, indicative of a motor neuron deficit. **(D)** Both transgenic lines had shorter survival times compared to non-transgenic (NT) controls (NT vs. WT:  $p = 0.0005$ , NT vs. G298S:  $p = 0.0035$ ), but no differences were found between the two transgenic lines (WT vs. G298S:  $p = 0.1260$ ). **(E-G)** RNA-Seq analysis on isolated mouse spinal cords revealed several examples of exon skipping (arrows) in both transgenic lines. **(H)** We further validated these findings by transfecting mouse N2a cells with human TDP-43 and performing RT-PCR to test whether TDP-43 expression alone was sufficient to induce exon skipping



**Fig. 2** (See legend on next page.)

(See figure on previous page.)

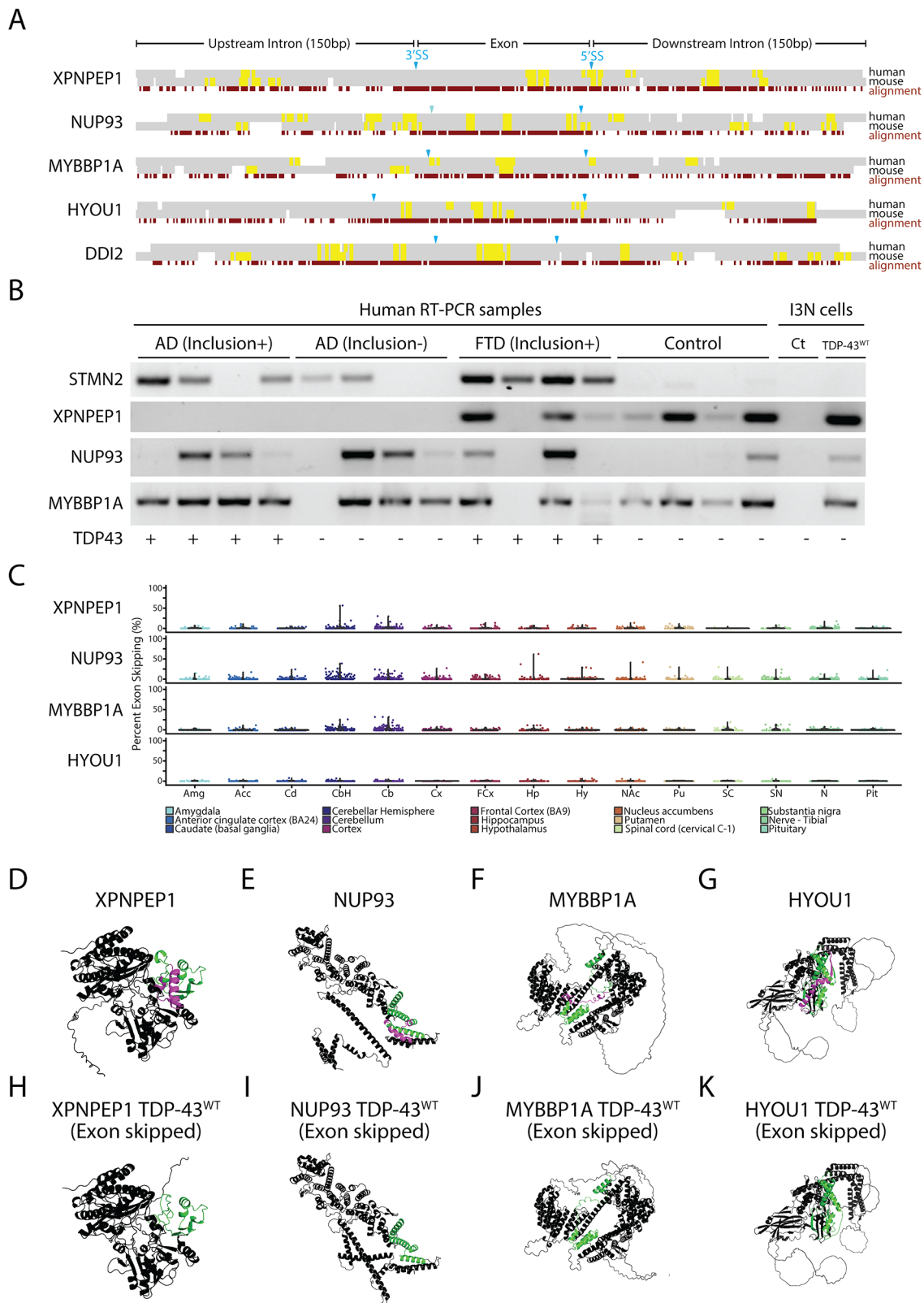
**Fig. 2** TDP-43 overexpression induces exon repression in humans. **(A)** Human i3Neurons were transduced with lentivirus expressing human TDP-43 and sequenced to compare exon skipping between mouse and human. **(B)** RNA-Seq analysis revealed numerous genes with skipped exons that are involved in a variety of molecular pathways related to intellectual disability, synaptic activity, and mitochondrial proteins **(C)**. We identified exons with particularly high levels of exon skipping in the genes *HYOU1*, *NUP93*, and *XPNPEP1* (arrows). **(D-F)** When cross-referenced with datasets from transgenic mice, we found that exons repressed in humans were not repressed in mice. Using RT-PCR, we validated skipping events in i3Neurons with primers located in exons adjacent to the repressed exon **(G)** or primers that spanned a skipped junction **(H)**. **(I)** Analysis of UG repeats in human skiptic exons revealed that TDP-43's consensus motifs are found in both the repressed exon itself and adjacent intronic sequences. UG motifs appear slightly more frequently around the downstream 5' splice site, but with far shorter UG repeat lengths than those found adjacent to cryptic exons [73]

with littermate controls (Fig. 1C) and animal weights for both transgenic mouse lines markedly decreased over time (Supplementary Fig. 1). The Kaplan-Meier survival curve indicated that lifespans of individuals from both transgenic lines were significantly shorter than their littermate controls (Fig. 1D). Pathological examination showed that TDP-43 overexpression was confined to the nucleus of neurons from the cortex and spinal cord, without any cytoplasmic inclusions in either transgenic line. Furthermore, histological analysis revealed reduction in axonal diameter, neuromuscular denervation and muscle degeneration in both transgenic lines, reflecting the severity of their condition (Supplementary Fig. 1). In summary, both TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> mouse lines exhibited similar phenotypes of mild motor deficits that appeared to be mutation-independent.

We next hypothesized that the observed phenotypes could be explained by changes in gene expression or alternative splicing due to the relatively low levels of TDP-43 overexpression. To profile these transcriptional changes, we surgically isolated the ventral horn of the spinal cord and conducted bulk RNA sequencing (RNA-Seq) on the TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> lines, as well as non-transgenic controls. TDP-43 has been well-described as a splicing repressor and loss of TDP-43 function leads to the inclusion of nonconserved cryptic exons. Interestingly, analysis of transgenic RNA-Seq data revealed that overexpression of both TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> leads to the skipping of constitutive conserved exons (Fig. 1E-G). To verify that these splicing events were indeed due to TDP-43 overexpression and not due to the secondary effects of neuronal degeneration, we overexpressed TDP-43 in mouse N2a cells and performed a RT-PCR to detect the same targets identified from TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> transgenic RNA-Seq data. As predicted, we observed prominent constitutive exon skipping only following TDP-43 overexpression (Fig. 1H). Together, these data indicate that even mild overexpression of TDP-43 (50% increase over controls) is sufficient to induce aberrant exon skipping and that exon skipping could be mediating the dose-dependent toxicity of TDP-43 overexpression.

Given the evolutionary conservation of the coding exons skipped by TDP-43 overexpression in mice, we hypothesized that syntenic exons in the human genome might also be skipped by TDP-43 overexpression in

human cells. To profile exon skipping, we transduced human i3Neurons [64] with lentivirus expressing TDP-43 and examined the splicing patterns of the equivalent exons observed to be skipped in mice (Fig. 2A). Unexpectedly, a large majority of skipped exons in mice were not skipped in human cells following TDP-43 overexpression, despite a ~1.6x fold increase in TDP-43 protein (Supplementary Fig. 2). To understand these differences, we expanded our RNA-Seq analysis on transduced i3Neurons to include all splicing events and found that exon skipping still occurred across multiple genes (78 skipped exons, Supplementary Fig. 3) but at sites entirely different from those found in mice (Supplementary Table 1). We only found two exceptions in *DDI2* and *SLC6A6* where skipping occurred for equivalent mouse and human exons (Supplementary Fig. 3). Importantly, *SLC6A6* has already been validated as a skiptic event in the TDP-43<sup>M323K</sup> mouse line and human fibroblasts derived from ALS patients carrying *TARDBP* pathogenic mutations [56]. The genes affected by exon skipping regulate a variety of cellular pathways including those associated with intellectual disability, synaptic activity, and mitochondrial proteins (Fig. 2B-C). Three of the clearest splicing repression events (~90% reduction) include exons in *HYOU1*, *NUP93*, and *XPNPEP1*, where the corresponding exon in mice remains constitutively spliced (Fig. 2D-F). We confirmed that these human skipped exons we evolutionarily conserved (Supplementary Fig. 4) and further validated the RNA-Seq results by RT-PCR analysis (Fig. 2G-H). UG repeats serve as the consensus binding site for TDP-43. An analysis of exons repressed by TDP-43 overexpression reveals the presence of short UG motifs that may be responsible for species-specific exon skipping, but these UG motifs are not as long as UG repeats associated with cryptic exons (Fig. 2I, Supplementary Fig. 5) [15, 16, 59, 65]. This suggests that smaller UG repeats may be available for TDP-43 to bind when protein concentrations of nuclear TDP-43 exceed a certain threshold, as indicated by a limited number of eCLIP motifs adjacent to skipped exons (Supplementary Fig. 6). Overall, our data indicate that while TDP-43 overexpression leads to aberrant exon skipping across the transcriptome, the exons that are skipped appear to be species-specific and generally exhibit shorter UG repeats than cryptic exons (Fig. 3A).



**Fig. 3** (See legend on next page.)

(See figure on previous page.)

**Fig. 3** TDP-43 exon skipping events are found in aging human brains but do not correlate with disease. **(A)** Alignment of syntenic mouse (mm10) and human (hg38) genomic sequences surrounding exons repressed by TDP-43 overexpression in human cells. Constitutively spliced exons in the genes *XPNPEP1*, *NUP93*, *MYBBP1A*, and *HYOU1* are skipped in human cells but not mouse cells when TDP-43 is overexpressed. By contrast, the exon in *DDI2* is repressed in both mouse and human cells. UG motifs (yellow highlights) are slightly enriched around the 3' and 5' splice sites (Supplementary Fig. 5). **(B)** We performed RT-PCR to amplify cryptic junctions or exon-repressed junctions in human brain samples from patients with AD pathology with or without TDP-43 inclusions, frontotemporal dementia with inclusions, and control patients who did not have TDP-43 inclusions and profiled the *STMN2* cryptic exon [16, 17]. RT-PCR analysis showed that exon skipping occurred in both control and disease samples. **(C)** Exploration of skipping events through normal human aging was performed by analyzing RNA-Seq datasets from the Genotype-Tissue Expression (GTEx) project. PSI values of skiptic exons across human brain regions are shown for the age range of 60 to 69 years old (all age ranges are available in Supplementary Fig. 7). Exon skipping is found at low levels in most of the different brain areas analyzed, with slightly higher levels in the cerebellum. **(D-K)** AlphaFold2 was used to model protein structures with **(D-G)** and without **(H-K)** skipped exons. Purple highlights indicate repressed exons while green highlights indicate flanking amino acid sequences

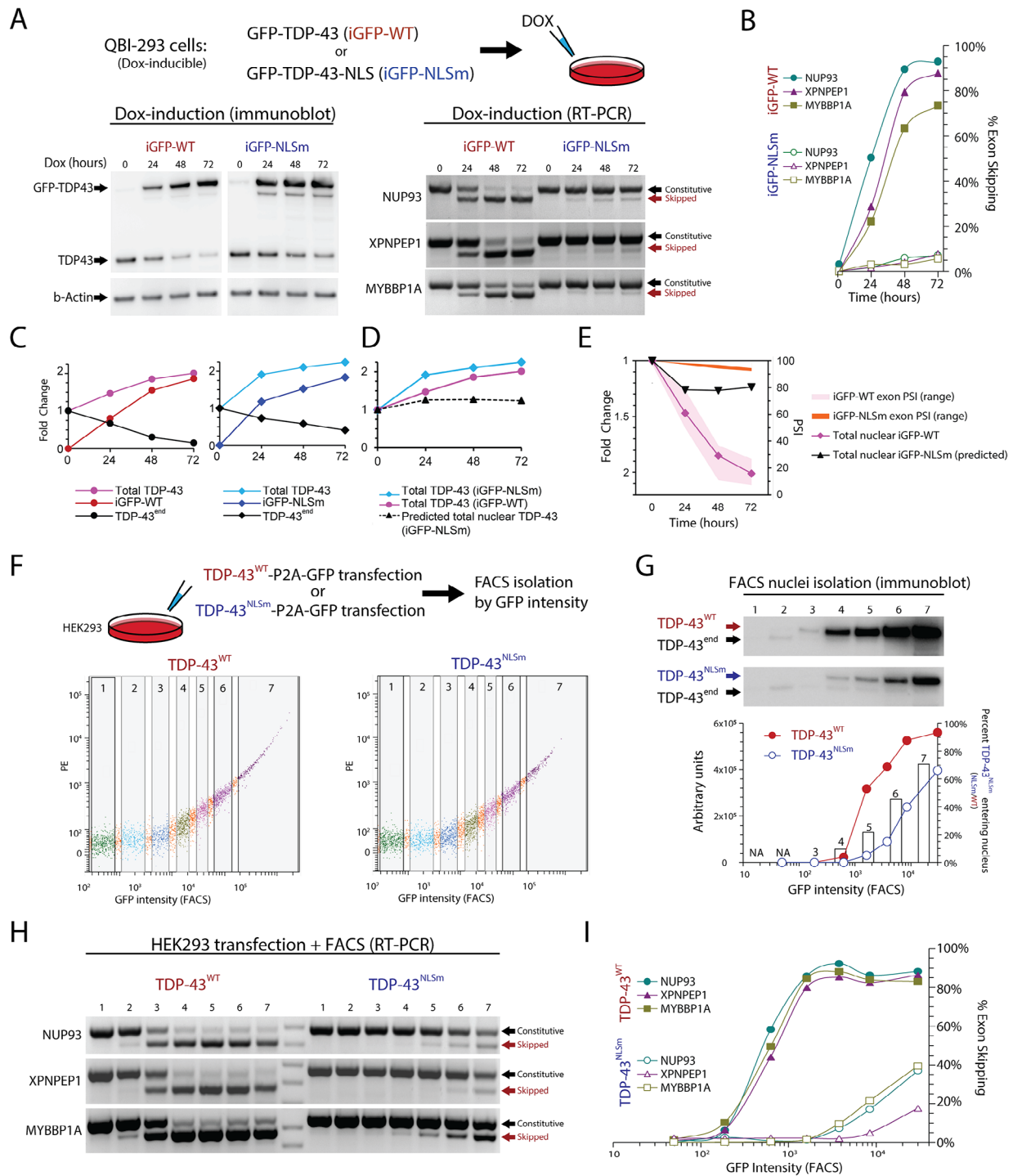
Next, we wanted to explore whether TDP-43-mediated exon skipping represents a potential mechanism underlying the pathogenesis of ALS and TDP-43-related dementias. To study this, we designed single-band RT-PCR primers that could selectively amplify the unique junction produced by constitutive exon skipping. We profiled brain samples from FTD with TDP-43 inclusions, AD with or without TDP-43 inclusions, TDP-43 negative human brain controls, and i3Neurons. Previous studies have demonstrated that a cryptic exon found in the gene *STMN2* is a robust biomarker for ALS-FTD [17–19, 66]. As expected, the *STMN2* cryptic exon was detected in all four FTD samples and in 3 out of 4 AD TDP-43 positive samples and 2 out of 4 AD TDP-43 negative samples (Fig. 3B). AD cases without TDP-43 inclusions may still exhibit cryptic exons (*STMN2*) due to the observation of TDP-43 nuclear clearance without cytoplasmic aggregation [24, 67, 68]. However, skipping events in at least three genes (*HYOU1*, *NUP93*, and *XPNPEP1*) could be detected in both disease cases and controls.

Since exon skipping was detected in some control cases, we explored whether aberrant exon skipping could be correlated with aging. To study this, we explored publicly available RNA-Seq archives using Snaptron and Recount3 [69–72], queried each skipped exon across thousands of brain samples available through the GTEx repository [73], and generally observed low levels of exon skipping, with slightly higher skipping frequencies in cerebellum tissue (Fig. 3C, Supplementary Fig. 7). Our data suggest that aberrant exon skipping, induced by excessive levels of nuclear TDP-43, is unlikely to play a significant role in disease pathogenesis.

Since aberrant exon skipping can occur independent from disease status, it will be important to avoid exon skipping when modeling TDP-43 proteinopathy. Overexpression of TDP-43 can be a valuable tool for investigating cytoplasmic toxicity, gain-of-function mechanisms, and post-translational modifications that may impact TDP-43 binding and other pathological interactions [74–77]. Indeed, certain methods that induce cytoplasmic aggregation of TDP-43 can mimic loss-of-function [78, 79]. However, dose-dependent toxicity resulting from exon skipping may confound results when analyzing

phenotypes from overexpression model systems, particularly since TDP-43 induced exon skipping appears to be highly species specific. Approximately 43% (30/69, Supplementary Table 1) of skipped exons are predicted to induce NMD, although skipped exons that retain reading frame are still predicted to impact protein structure (Fig. 3, Supplementary Fig. 8). Using the AlphaFold 2 protein prediction artificial intelligence system [80], we investigated the consequences on protein structure due to exon skipping and found significant impacts on protein structure and folding. For example, the *XPNPEP1* isoform that results from exon skipping leads to complete disruption of the second domain of the protein (Fig. 3D, H). For *HYOU1*, we observe a loss of alpha helical domains in the C-terminal region of the protein (Fig. 3G, K). In *NUP93*, we predict that exon skipping would lead to dramatic changes in the C-terminal region that is crucial for nuclear pore complex assembly (Fig. 3E, I). These results suggest that inframe exon skipping can still dramatically affect protein structure and function, thereby leading to cellular toxicity.

A promising alternative to overexpression of wildtype TDP-43 (TDP-43<sup>WT</sup>) that could avoid exon skipping is the overexpression of TDP-43<sup>NLSm</sup>, where the nuclear localization signal is mutated [50, 52, 58]. Transgenic mice (rNLS8) have been generated where expression of TDP-43<sup>NLSm</sup> can be induced by doxycycline (Dox) under the control of a neuron-specific driver line, the human NEFH-tTA promoter [53]. However, since TDP-43<sup>NLSm</sup> can enter the nucleus passively [54, 81], we wanted to determine the maximum level of TDP-43<sup>NLSm</sup> expression at which exon skipping can be detected. For this we used previously established doxycycline (Dox) inducible HEK-293 stable cell lines [58] to conditionally express TDP-43<sup>WT</sup> (iGFP-WT) or TDP-43<sup>NLSm</sup> (iGFP-NLSm) (Fig. 4A). Using RT-PCR, we found that exon skipping was present following doxycycline-dependent induction of either TDP43<sup>WT</sup> or TDP43<sup>NLSm</sup> Dox induction. However, the percentage of exon skipping for iGFP-NLSm induction reached only approximately 9% while iGFP-WT induction led to exon skipping of ~90% (Fig. 4B). As previously described, induction of either iGFP-WT or iGFP-NLSm causes autoregulation of endogenous



**Fig. 4** Overexpression of the  $\Delta$ NLS mutant TDP-43 (TDP-43<sup>NLSm</sup>) induces exon repression when expressed at higher levels than wildtype TDP-43. QBI-293 cells with a dox-inducible cassette expressing TDP-43-GFP (iGFP-WT) or TDP-43 with a mutated NLS (iGFP-NLSm) were exposed to Dox for 0, 24, 48, 72 h in triplicate. **(A)** Immunoblot of TDP-43 levels after Dox induction (left) and RT-PCR for *NUP93*, *XPNPEP1* and *MYBBP1A*. **(B)** RT-PCR quantification of iGFP-NLSm, exon repression reached ~9% compared to ~95% in iGFP-WT induction. **(C)** Total TDP-43 protein levels for WT and NLSm reached ~2 times normal. **(D, E)** Estimated nuclear TDP-43 levels based on TDP-43<sup>NLSm</sup> passive diffusion. **(F, G)** HEK-293 cells transfected with TDP-43<sup>WT</sup>-P2A-GFP or TDP-43<sup>NLSm</sup>-P2A-GFP were FACS sorted into seven fractions by GFP intensity. Immunoblotting of isolated nuclei showed increasing TDP-43<sup>NLSm</sup>/TDP-43<sup>WT</sup> ratio with higher expression. **(H)** RT-PCR showed progressive exon skipping with TDP-43<sup>WT</sup> or TDP-43<sup>NLSm</sup> overexpression, performed in duplicate. **(I)** With strong expression, TDP-43<sup>WT</sup> can repress exons by ~95% but TDP-43<sup>NLSm</sup> only by ~40%. Using these data, we estimate the proportion of TDP-43<sup>NLSm</sup> in the Dox-inducible system, dotted line (D). Levels of total TDP-43<sup>WT</sup> and total predicted TDP-43<sup>NLSm</sup> in the nucleus are plotted together with their respective exon expression levels (E)



TDP-43, but also increases the total amount of TDP-43 protein (Fig. 4C-D). We estimate that a 50% increase in total TDP-43 compared to its endogenous level is enough to induce at least 40% of exon skipping. By comparison, a 100% increase in total TDP-43, due to iGFP-NLSm induction, leads to only 9% exon skipping (Fig. 4E). If we assume that 45% of iGFP-NLSm can passively diffuse into the nucleus (Table 1), we can estimate the total TDP-43 with the following calculation for Fig. 4D and E: [Total TDP-43] = [endogenous TDP-43] + 0.45 \* [iGFP-NLSm].

In animal models, first month of doxycycline induction results in overexpression of about 8 to 10-fold total TDP-43 (TDP-43<sup>NLSm</sup> + endogenous TDP-43), suggesting that excessive levels of TDP-43<sup>NLSm</sup> may induce higher levels of exon skipping [58]. To test whether higher levels of TDP-43<sup>NLSm</sup> overexpression promotes higher passive diffusion into the nucleus, we transfected HEK-293 cells with TDP-43<sup>WT</sup>-P2A-GFP or TDP-43<sup>NLSm</sup>-P2A-GFP and used FACS to obtain seven population fractions based on GFP intensity (Fig. 4F). We performed nuclei isolation on each fraction and quantified TDP-43 levels by immunoblotting and found that endogenous TDP-43 (TDP-43<sup>end</sup>) was eliminated by autoregulation from TDP-43<sup>WT</sup> transfection in FACS populations with higher GFP intensity. This autoregulation was less evident in TDP-43<sup>NLSm</sup>, as nuclear TDP-43<sup>NLSm</sup> did not reach levels that were sufficient to completely repress TDP-43<sup>end</sup> even in the highest GFP intensity population. By comparing the ratio of TDP-43<sup>WT</sup>/TDP-43<sup>NLSm</sup> at equivalent FACS populations, we estimated that 10–45% of TDP-43<sup>NLSm</sup> can enter the nucleus. The wide range in diffusion percentages reflects the non-linear increase in TDP-43<sup>WT</sup> transfection, as total TDP-43 levels begin to saturate and plateau despite an increase in GFP fluorescence (Fig. 4G). To estimate the minimal level of TDP-43 required to induce exon

skipping, we also isolated RNA from each of the seven FACS populations isolated by GFP intensity and evaluated exon skipping using 2-band RT-PCR (Fig. 4H). We found that TDP-43<sup>WT</sup> overexpression induced 90% exon skipping while TDP-43<sup>NLSm</sup> saturated at 40% exon skipping (Fig. 4H). Since TDP-43<sup>NLSm</sup> induces exon skipping in FACS fraction six at equivalent levels as the iGFP-NLSm Dox inducible system, i.e. both induce 9% exon skipping, we can use immunoblot analysis of isolated nuclei to estimate that nuclear TDP-43<sup>NLSm</sup> is approximately 40% of nuclear TDP-43<sup>WT</sup> (Fig. 4G). With these estimates, we extrapolated that total TDP-43 levels 1.1 to 1.5-fold above normal can begin to induce exon skipping. Together this suggests that cell line and animal models using TDP-43<sup>NLSm</sup> could cause exon skipping toxicity if expression reaches sufficiently high levels. Estimations on the levels of TDP-43 required to induce exon skipping and the diffusion of TDP-43<sup>NLSm</sup> are summarized in Table 1.

## Discussion

In this study, we explore the molecular mechanisms by which excessive levels of TDP-43 in the nucleus can induce cellular toxicity. Our findings reveal that TDP-43 overexpression leads to skipping of constitutive exons that are normally incorporated into mRNAs under steady state conditions. This aberrant exon skipping appears to be driven by high concentrations of nuclear TDP-43 binding to suboptimal UG motifs located near these constitutive exons, although further studies are necessary to understand the exact mechanism of splicing repression.

Indeed, we observe that nearly all exons skipped due to TDP-43 overexpression are species-specific, at least when comparing between TDP-43 overexpression in mouse and human neurons. It is important to note that there is

Excess TDP-43 required to induce exon skipping	
Fold change increase nuclear TDP-43 over baseline endogenous	Exon skipping
1.1-1.5x	Minimum exon skipping
1.5-2x	50-100% exon skipping
>2x	Maximal exon skipping

Estimated diffusion of TDP-43 <sup>NLSm</sup> ( $\Delta$ NLS TDP-43) into the nucleus	
Experimental condition	Percentage of total TDP-43 <sup>NLSm</sup>
Immunoblot of FACS isolated nuclei from transfected cells	10-45%
Immunoblot of inducible iGFP-WT and iGFP-NLSm cells	45%
Fold change of GFP fluorescence intensity by FACS	2-5%*

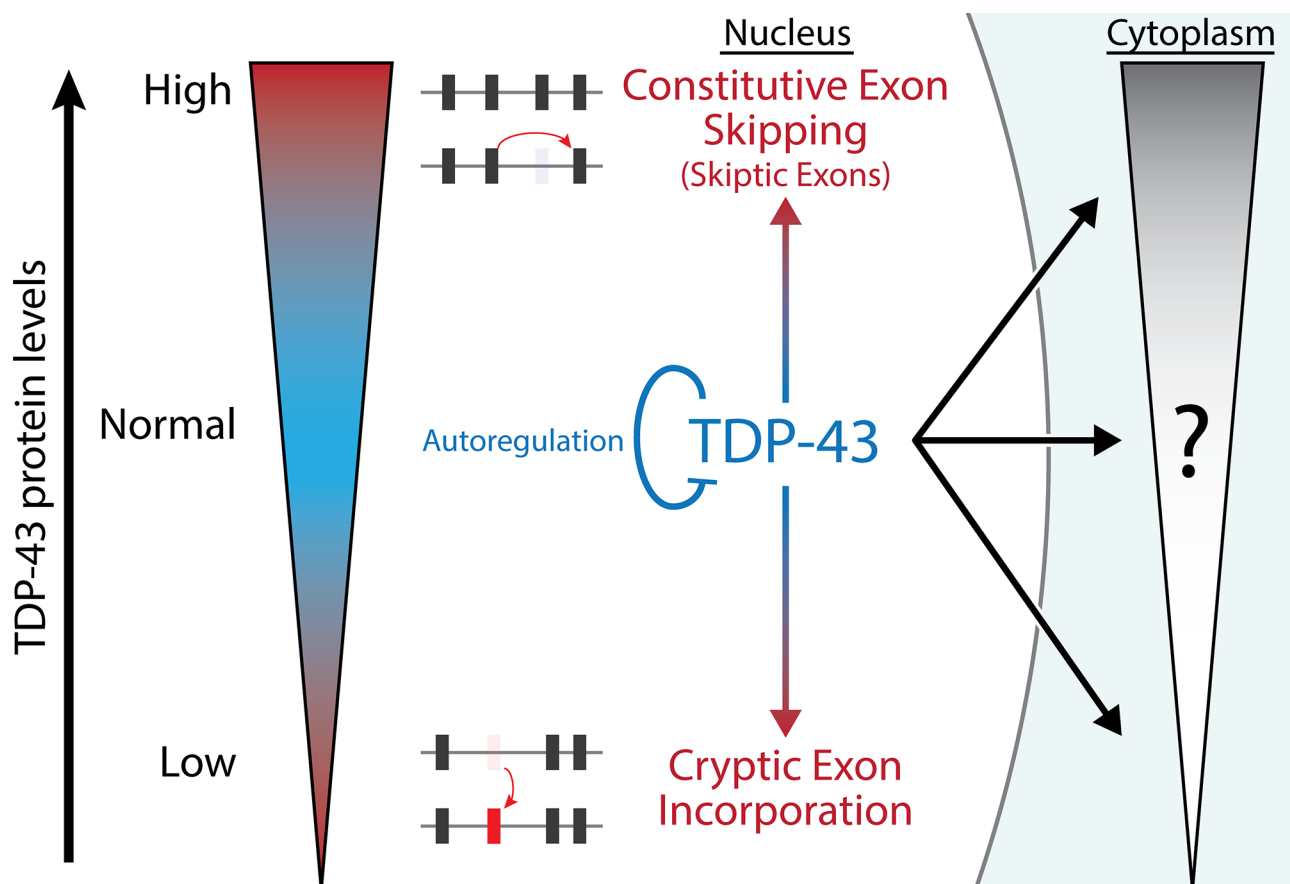
\*GFP is indirect measurement and may not accurately estimate the amount of nuclear  $\Delta$ NLS-TDP-43

**Table 1** Estimates for the amount of excess nuclear TDP-43 that is required to induce exon skipping across different experimental conditions. Our results suggest that an increase in TDP-43 that is greater than 1.1 to 1.5 fold over normal protein levels may lead to exon skipping toxicity. Models for TDP-43 overexpression should avoid increasing nuclear TDP-43 and aim to strictly limit excess TDP-43 to only the cytoplasm

no clear consensus on the precise definitions and criteria used to classify constitutive, cryptic, and skiptic exons within the TDP-43 field. In this study, we consider exons with high percent spliced-in (PSI) values (>90%) under steady-state conditions as constitutive, acknowledging that this threshold may differ from prior studies, and define skiptic exons as exons with high PSI values that are aberrantly skipped due to TDP-43 gain-of-function. We recognize that some identified skiptic events may represent increased alternative splicing rather than skipping of strictly constitutive exons, but further community effort is needed to precisely define these categories. Previous work has identified constitutive exon skipping (skiptic exons) in a knock-in mouse model that carries a TDP-43 mutation in its low complexity domain (Tardbp<sup>Q331K</sup>) and suggested that skiptic splicing is a gain-of-function directly associated with disease-associated mutations [56]. Our data suggests that the observed increase in total TDP-43 in heterozygous and homozygous Tardbp<sup>Q331K</sup> mice likely mediates skiptic splicing. By contrast, some disease-associated TDP-43 mutations appear to cause

loss-of-function that can result in the incorporation of cryptic exons [82–84]. It remains to be seen whether other mutations in TDP-43's C-terminal domain can also increase total TDP-43 levels and lead to constitutive exon skipping. In a limited comparison of PSI values between the TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> transgenic models of this study, we do not detect skiptic events that are unique to TDP-43<sup>G298S</sup> (Supplementary Fig. 9), but this does not rule out the possibility that other variants having mutation-specific effects on splicing [17, 55] (Supplementary Data Files 3–5). Knock-in models of TDP-43 mutations may recapitulate potential shared mechanisms of disease between mouse and human, but our study suggests that skiptic exons and affected cellular pathways will be largely species specific.

Furthermore, while skiptic splicing could be detected in some human brain samples, these aberrant splicing events could also be detected in non-disease controls. By contrast, cryptic exons were exclusively detected in human disease samples, indicating their potential significance in disease pathogenesis. These findings have



**Fig. 5** Summary Diagram. TDP-43 is a highly autoregulated protein due to different forms of cellular toxicity when TDP-43 protein levels are either too low (cryptic exon incorporation) or too high (constitutive exon skipping, i.e. skiptic exons). Our study has demonstrated that these splicing deficits are linked specifically to nuclear TDP-43, whereas toxicity due to cytoplasmic TDP-43 remains to be fully elucidated. Future studies that avoid constitutive exon skipping (skiptic exons) may identify biomarkers for cytoplasmic-specific TDP-43 toxicity

important implications for the interpretation and use of TDP-43 overexpression models in neurodegenerative disease research.

Our study underscores the need for caution when interpreting data obtained from TDP-43 overexpression models. We believe that TDP-43 induced exon skipping can be avoided when generating such models by ensuring that overexpressed TDP-43 is restricted to the cytoplasm and limiting nuclear TDP-43 protein levels to 130–150% above steady state. Cytoplasmic gain of function and nuclear loss of function of TDP-43 may both contribute to the pathogenesis of neurodegenerative diseases, but skiptic splicing may confound the interpretation of some model systems. For example, we observe minimal exon skipping in cell lines that stably express  $\Delta$ NLS-TDP-43, but rNLS8 mice exhibit levels of nuclear TDP-43 that may be sufficient for skiptic splicing. Likewise, it remains to be determined whether AAV-mediated delivery of TDP-43 can also induce exon skipping, or whether TDP-43 autoregulation can minimize excessive nuclear TDP-43.

A notable limitation of our study is the lack of direct evidence connecting exon skipping with cell survival and disease pathology *in vivo*. We describe a potential pathogenic mechanism by which overexpression of various forms of TDP-43 (wildtype, mutant, or  $\Delta$ NLS) induce aberrant exon skipping, but the context in which these exon skipping events occurs has not been fully addressed. TDP-43 overexpression may have other unexplored off-target effects, with exon skipping perhaps representing only one part of a broader impact. Further research is needed to assess the consequences of exon skipping on cell viability and disease progression, and to explore the full range of TDP-43's effect in the cytoplasm. For example, TDP-43 overexpression has recently been linked to the misregulation and accumulation of NPTX2 [85]. Although significant reductions in skiptic exons were not observed, exon skipping was nevertheless detected in bulk sequencing data at a range of % to % upon induction with TDP-43 (Supplementary Fig. 10), suggesting the possibility that a subset of cells may exhibit higher levels of exon skipping.

In conclusion, our study provides new insights into the complex molecular mechanisms underlying TDP-43 gain- and loss-of-function models. Our findings suggest that TDP-43 autoregulation is a highly conserved mechanism because both reduction and increases in TDP-43 protein levels lead to toxicity; depletion of nuclear TDP-43 leads to cryptic exon incorporation, while excess nuclear TDP-43 leads to constitutive exon skipping (Fig. 5). Future models of TDP-43 pathology should consider both cryptic exon inclusion and constitutive exon skipping to capture the complex role of TDP-43 in neurodegenerative diseases.

## Methods

### Ethics declarations

The authors declare no competing interests.

### Antibodies

Antibody (Company, Catalog #)	Method (Dilution)
Rabbit anti-Mouse TDP-43 N-Term (Proteintech, 10782-2)	HEK-293 cells immunoprecipitation, mouse tissues immunohistochemistry, mouse cortex and spinal cord immunohistochemistry, i3Neuron stable cell line immunofluorescence
Rabbit anti-Mouse TDP-43 C-Term (Proteintech, 12892-1)	mouse tissues immunoprecipitation, mouse cortex and spinal cord immunohistochemistry
Rabbit anti-Synaptophysin (Life Technologies, Z66)	mouse tissues immunohistochemistry
Mouse anti-Pan-Axonal neurofilament (Covance, SMI-312R)	mouse tissues immunohistochemistry
Mouse anti-Human TDP-43 (Novus, 2E2-D3)	mouse cortex and spinal cord immunohistochemistry
Mouse anti-GAPDH (Sigma, GAPDH-71.1)	mouse cortex and spinal cord immunohistochemistry
Mouse anti-FLAG (Sigma, F1804)	i3Neuron stable cell line immunofluorescence
Rabbit anti-GAPDH (Cell Signalling, 2118)	i3Neuron stable cell line immunofluorescence
Mouse anti-Tubulin (ThermoFisher, 14-4502-82)	i3Neuron stable cell line immunofluorescence
Rabbit anti-MYBBP1A (Proteintech, 14524-1-AP)	i3Neuron stable cell line immunofluorescence
Rabbit anti-NUP93 (Proteintech, 13077-1-AP)	i3Neuron stable cell line immunofluorescence
Mouse anti-XPNPEP1 (Abcam, ab123929)	i3Neuron stable cell line immunofluorescence
Mouse anti-TDP-43 (Abcam, ab104223)	QBI-293 stable cell line immunofluorescence
Rabbit anti-GFP (Cell Signalling, mAb 2956)	QBI-293 stable cell line immunofluorescence
Mouse anti-Beta Actin (Proteintech, 66009-1)	QBI-293 stable cell line immunofluorescence

### Transgenic mouse generation

A Thy1.2 expression cassette on a pUC18 backbone was used to generate transgenic mice. Plasmids pTSC-TDP-WT and pTSC-TDP-G298S were constructed and submitted to the Transgenic Core Laboratory at the Johns Hopkins University School of Medicine for pronuclear injection using the hybrid mouse strain C57B6;SJL. Potential founders were screened by tail cutting, genomic DNA extraction, and PCR. The WT and G298S lines were maintained in the hybrid background and phenotypic characterization was started at the F3 generation. All mouse experiments were approved by the Johns Hopkins University Animal Care and Use Committee.

### Hanging wire tests

F3 generation was followed to the end-stage (death or inability to right from a lateral decubitus position), and survival data was recorded. A subset of this population was selected for monthly hanging wire tests. Briefly, a mouse was placed on a wire grid, and the grid was shaken to encourage the mouse to grip the wires. The grid was then turned upside down and held level while the time for the mouse to fall was recorded, the endpoint was 60 s. Each mouse was tested 3 times, with the maximum hang time recorded as the result.

### Tissue preparation for sectioning

Mice were anesthetized with an intraperitoneal injection of 15% chloral hydrate and anesthesia was monitored with limb and corneal reflex checks. Portions of quadriceps muscles were snap-frozen for sectioning. Mice were transcardially perfused with 50–100 mL of cold PBS, and then fixed on 4% PFA. The L3 and L4 dorsal root ganglia and attached dorsal and ventral roots were identified, removed, placed in a fixation buffer with 2% glutaraldehyde overnight, washed in PBS and embedded in Epon for EM. The brain and spinal cord were dissected and separated into right/left halves. One half of each was placed in a fixation buffer overnight at 4 °C with gentle agitation, and then held in PBS prior to embedding in paraffin and sectioning. The other half brain was half spinal cord was placed in a fixation buffer for 2 h, and then switched to sterile PBS with 30% sucrose overnight at 4 °C to be embedded in O.C.T and frozen in isopentane for cryostat sectioning.

### Immunofluorescence and immunohistochemistry

Paraffin-embedded brain samples were sectioned in the sagittal axis and spinal cord samples were sectioned in the cross-sectional axis onto slides. Slides used for immunohistochemistry were incubated at 60 °C for 30 min, then deparaffinized in xylene and ethanol. Antigen retrieval was accomplished by incubating slides for 5 min in boiling 10mM sodium citrate buffer (pH 6.0). Slides were washed with PBS and blocked and stained using appropriate primary antibodies and reagents from the Vectastain Elite ABC kit (Vector Labs). Diaminobenzidine exposure was titrated to optimal contrast, sections were counterstained with Mayer's hemalaun, and then slides were dehydrated using ethanol and xylene and mounted.

Frozen gastrocnemius muscle was cut into 40 µm longitudinal sections using a freezing sliding microtome (Leica). Sections were separated into wells on a 12-well plate and blocked in IF blocking buffer: PBS with 5% normal goat serum and 0.5% Triton-X, slides were incubated overnight at 4 °C on primary antibody (rabbit anti-Synaptophysin and mouse anti-SMI-312), diluted in blocking solution. Sections were then washed with PBS with 0.5% Triton-X, incubated with  $\alpha$ -bungarotoxin and goat anti-rabbit and anti-mouse secondary antibodies conjugated to Alexa Fluor-488 (Invitrogen), and washed again. Sections were spread on slides and coverslips were attached using Prolong Gold Antifade Reagent (Life Technologies). Slides were examined using a Zeiss LSM 510 Meta confocal microscope.

### Muscle histology

Frozen quadriceps muscle embedded in O.C.T. were cut into 10 µm cross-sections onto slides using a cryostat

(Leica). For H&E staining of muscle, sections were covered for 20 min with a fixation buffer (4% PFA in PBS). Sections were rinsed with distilled water, stained with Mayer's Hemalaun, rinsed, and stained with Eosin. After a final rinse, samples were dehydrated in ethanol and xylene, and coverslips were applied. For esterase staining, 25%  $\alpha$ -naphthyl acetate, 5% acetone, 0.1% Pararosaniline HCL, and 0.1% Sodium Nitrate in 0.2 M Sodium Phosphate solution was used. The solution was applied to quadriceps sections for 5 min, rinsed in running tap water for several minutes and slides were then dehydrated in ethanol and xylene, mounted. Esterase and H&E-stained sections were analyzed under light microscopy.

### i3neurons TDP-43 lentivirus transduction

i3Neurons [86] were transduced with lentivirus containing N-terminal Flag-tagged wild type TDP-43 at 1MOI, 2MOI, and 4MOI, respectively, at Brain physiological stage day 11. Neurons were harvested on Day 14 and dried ice frozen. Total RNA was isolated using Trizol extraction and used for downstream RNA sequencing analyses.

### TDP-43 dox-inducible QBI-293 stable cell lines

iGFP-WT and iGFP-NLSm inducible cell lines were kindly provided by Silvia Porta and Virginia Lee [52] and cultured on Dulbecco's Modified Eagle Medium supplemented with 10% FBS (Corning, 35-010-CV), and 1% Penicillin-Streptomycin (ThermoFisher Scientific, 15,070,063), L-glutamine (20 mM, Corning Cellgro, Manassas, VA) with G418 (400 µg ml<sup>-1</sup>, Calbiochem, La Jolla, CA). Cells were induced with Dox ug/mL and collected after 24 h, 48–72 h of induction.

### HEK-293 cell culture, transfection, FACS separation and nuclei isolation

HEK-293 cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% FBS (Corning, 35-010-CV), 1x GlutaMAX (ThermoFisher Scientific, 35,050,061), and 1% Penicillin-Streptomycin (ThermoFisher Scientific, 15,070,063). For overexpression of TDP-43<sup>WT</sup> and TDP-43<sup>NLS</sup> ORF expression cassettes were cloned into pAAV-CBh-mKate2-IRES-MCS (a gift from Marcella Patrick, Addgene plasmid # 105,921) and transfected on HEK-293 cells with Lipofectamine 3000 (Thermo Fisher Scientific, L3000-008) following the manufacturer's protocol. Two days of transfection single cell suspensions were obtained using TrypLE (ThermoFisher Scientific, 12,604,013) and sorted by GFP fluorescence intensity on a BD FACSCalibur in the JHMI Ross Flow Cytometry Core Facility. Nuclei were isolated following the 10x Genomics nuclei isolation protocol from cell suspensions.

### Immunoblotting

Tissues and cells were digested in RIPA Buffer with 1%, protease inhibitor (Roche Complete ULTRA mini tablet+EDTA), and phosphatase inhibitor (Roche PhosStop). Samples were centrifuged, and supernatants were saved. Protein concentrations in the supernatants were determined using a BCA assay (Pierce), and 20 µg of total protein was loaded into each well of a 10–20% Tris-Glycine gel or a 4–12% Bis-Tris gel (Invitrogen). Protein was transferred to a PVDF membrane (Millipore), blocked with 5% BSA in TBS with 0.1% Tween-20. Membranes were incubated overnight at 4 °C with primary antibody diluted in blocking buffer, then washed three times with TBS with 0.1% Tween-20, and incubated for 2 h with secondary antibody (Goat anti-mouse IgG-HRP or Goat anti-rabbit IgG-HRP, Sigma) diluted in blocking buffer. Three more washes with TBS with 0.1% Tween-20 followed, and then membranes were soaked in ECL solution (EMD Millipore Immobilon), dried, imaged on a Bio-rad image. Densitometric analysis was performed using Quantity One software (Bio Rad).

### RNA extraction, library preparation, and RNA sequencing

Two ventral halves of spinal cords from TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> transgenic animals and one littermate control for each line were dissected and transferred immediately into RNAlater storage reagent (Life Technologies). Total RNA from spinal cord tissues was extracted using the RNeasy Mini Kit (Qiagen). Total RNA from human tissue, QBI-293 stable cell lines, and HEK-293 cells were extracted using Monarch Total RNA Miniprep Kit (New England BioLabs, T2010S). To prepare RNA-Seq libraries, the TruSeq Stranded Total RNA Library Prep Kit (Illumina) was employed. Subsequently, the sample libraries were sequenced on an Illumina HiSeq for spinal cord samples and NextSeq for cultured cells. The obtained data was transformed into FASTQ files after demultiplexing.

### RT-PCR and primers

Mouse animals were genotyped using the following primers: 5'-CGGAAGACGATGGGACGG TG, 5'-GCCAAACCCCTTTGAATGACCA, and 5'-AAGATG GCACGGAAGTCTAACCATG, was used to generate a 386 bp band (spanning TARDBP exons 2, 3, and 4) in transgenic mice only and a 241 bp internal control band in all mice. LunaScript RT SuperMix Kit was used for cDNA synthesis. The following primers were used for RT-PCR in mouse tissue: Ddi2-F1: CAGAGTGTGCT CGTTTGGCA, Ddi2-R1: GACTCGTCGGGCTACCA AC, product: 284 bp (Full) or 222 bp (Skip); Mrpl45-F1: AACTGTTTTCCGGACATGGT, Mrpl45-R1: TCG TACTCCTCCCAGGGTTT, product: 384 bp (Full) or 210 bp (Skip); Psm14-F1: GAGCCAGGTCCTTGT

GAGT, Psm14-R1: TTGGCTTGGAACACTGGATC A, product: 473 bp (Full) or 353 bp (Skip). The following primers were used for PCR in human cells: HYOU1-F2: TTCTATGACATGGGCTCAGGC, HYOU1-R2: ACT GCATCTCGGACGACAAA, product: 635 bp (Full) or 500 bp (Skip); NUP93-F2: GAGGCTGAAGAACATG GCAC, NUP93-R2: TCCCACAAAGCATGGCACTT, product: 496 bp (Full) or 411 bp (Skip); XPNPEP1-F2: GTTGGTGTGGACCCCTTGAT, XPNPEP1-R2: GAC CCACACCTTCTCCCTTG, product: 522 bp (Full) or 426 bp (Skip); MYBBP1A-F1: AAAGTCTGGGAGAGA AGCCC, MYBBP1A-R1: CGCAGAGCCTTCTCCTTCT G, product: 586 bp (Full) or 499 bp (Skip). The following primers were used for RT-PCR in human cells and brain samples: STMN2-F1: CTGCACATCCCTACAATGGC, STMN2-R1: CACAAGCCGCATTCACATTCA, product: 167 bp; HYOU1-F1: CCGTATGCACCATTGTGAC C, HYOU1-S5: CTGCTCAATCTCCTCCTGTGC, product: 290 bp; NUP93-S2: GGTCATATTGATAGAGCTTT TGATATCAGG, NUP93-R1: TGTACTGACAGTGTGC CGAC, product: 304 bp; XPNPEP1-S3: GCCTGGATTA CACAGGGCTATTT, XPNPEP1-R1: ATTCGGCTTCC AGACCCAAG, product: 177 bp; MYBBP1A-S2: CTGCA GCTAATTCTGGATGACAAG, MYBBP1A-R1: CGCAG AGCCTTCTCCTTCTG, product: 320 bp. Differences in RT-PCR product sizes were resolved on 4% agarose gels.

### Percent splice-in bioinformatic analysis

FASTQ files were aligned using STAR [87] and a filtered splice junction file was processed to calculate percent of splicing (PSI) values following methodology described in ASCOT [71, 72]. Briefly, FASTQ files were aligned to the mm10 (mouse) or hg38 (human) genomes to generate BAM files. BAM files were converted into bigWig files using Megadept [88]. GTF transcript annotations from GENCODE were used to calculate gene expression values and perform exon sequence analyses (GENCODE v45 was used for human, GENCODE M34 was used for mouse). We calculated the 5' PSI or 3' as the ratio between the number of mapped splice junctions from one exon to another over the total amount of splice junctions including that exon the 5' or 3' splicing site. PSI values between controls and TDP-43 overexpression were compared and visualized using UCSC genome browser [89]. Skiptic exon coordinates and PSI values are recorded in Supplementary Table 1. Custom code used in this project is available on Github: <https://github.com/rcarmen1/skipticexonscomparisons/tree/main>.

### Calculation of repeat frequency

Calculation of UG/GU repeat frequency around a +/- 400 bp window surrounding skiptic exons in Supplementary Fig. 5 was done by (i) masking all "UG" and "GU" as "YY"; (ii) replacing all "A," "C,"

“T,” “G” as “N”; (iii) identifying all pentamer and longer repeat sequences allowing for a single N insertion (i.e., “YYYYY,” “YYYYYY,” “YYYYYYY,” “YNYYYY,” “YYNYYY,” “YYYNYY,” “YYYYNY,” ...); (iv) assigning the “Y”s in sequences from step 3 a value of 1 and all other sequences a value of 0; (v) aligning all sequences to the 3’SS and 5’SS, applying a vertical summation, and then dividing by the total number of sequences to obtain the repeat frequency per nucleotide position upstream and downstream of the skiptic exon.

### Analysis of exon skipping in GTEx human samples

Splice junctions from the GTEx archive were queried using the Snaptron Web Services Interface and only brain, nerve and pituitary samples were collected [69, 73]. PSI values for skipped junctions were calculated as previously mentioned. We used the normalized quantification of transcripts per million (TPM) from the GTEx consortium (GTEx\_Analysis\_2017-06-05\_v8\_RSEMv1.3.0) to correlate PSI values and TDP-43 expression level [73]. Plots from different brain regions were made using RStudio.

### Protein structure modeling

Wildtype mRNA sequences for XPNPEP1 (ENST00000502935.6), NUP93 (ENST00000308159.10), MYBBP1A (ENST00000254718.9) and HYOU1 (ENST00000617285.5), were obtained from the GENCODE database and protein sequences were identified by translating the mRNA sequences with or without the exon of interest [90]. Wildtype structures for NUP93 and HYOU1 were downloaded from the AlphaFold Protein Structure Database. All other structures were generated using the Alpha Fold Monomer v2.0 pipeline (version 2.2.0) from the amino acid sequences on the Rockfish cluster at Johns Hopkins University [80, 91]. PyMOL was used to visualize the structures (Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.5.4, 2015).

### Statistical analysis

Mouse data was analyzed using Microsoft Excel and GraphPad Prism software, with P values calculated using unpaired Student’s t tests. P values less than 0.05 were considered significant. All quantitative values were reported as mean  $\pm$  standard deviation (SD).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13024-024-00732-w>.

Supplementary Material 1  
Supplementary Material 2  
Supplementary Material 3  
Supplementary Material 4

Supplementary Material 5  
Supplementary Material 6  
Supplementary Material 7  
Supplementary Material 8  
Supplementary Material 9  
Supplementary Material 10

### Acknowledgements

We thank X. Zhang and the Johns Hopkins Ross Flow Cytometry Core Facility. We also thank L. Orzolek, T. Creamer, and the Johns Hopkins Single Cell & Transcriptomics Core for sequencing services. We would like to thank Silvia Porta and Virginia Lee for kindly providing QBI-293 stable cell lines.

### Author contributions

J.P.L. and R.P.C.O. conceived and oversaw all aspects of the study. R.P.C.O. and J.P.L. conceived analysis of exon repression and validation in human cells and mouse lines. W.T. and P.C.W. oversaw TDP-43<sup>WT</sup> and TDP-43<sup>G298S</sup> mouse line development and behavioral and pathological analyses. Y.Y. and S.S. oversaw human iPSC neuronal culture studies. L.R.H. assisted with Dox inducible QBI-293 cell culture. K.C. and J.C.T. assisted with pathological characterization of human brain tissues. V.T., W.C. and K.B., assisted with data analysis and human RT-PCR. I.R.S. assisted with protein structure predictions. S.B., L.R.H., S.S., and P.C.W. provided valuable feedback on experimental design and manuscript writing. R.P.C.O. and J.P.L. drafted the manuscript. All authors approved the final manuscript.

### Funding

Part of this work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), supported by the National Science Foundation (NSF) grant number OAC 1920103. This work was supported by grants from the Alzheimer’s Association (J.P.L.), the Packard Center for ALS Research at Johns Hopkins to P.C.W., the ALS Association to P.C.W., the NIH (RF1MH123237 to S.B., P30AG066507 to J.C.T., U19AG033655 to J.C.T., R01NS123538 to L.R.H., RF1NS127925 to S.S., RF1NS113820 to S.S., R01AG078948 to S.S., R01NS095969 to P.C.W.) and the National Science Foundation Graduate Research Fellowship Program to I.R.S. under Grant No. DGE2139757.

### Data availability

RNA-Seq FASTQ files have been deposited in the NCBI Sequence Read Archive under BioProject accession number PRJNA1104758.

### Declarations

#### Ethics approval and consent to participate

Human brain samples were obtained following protocols authorized by the Institutional Review Board of the Johns Hopkins University School of Medicine. All animal procedures were in accordance strictly with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Johns Hopkins University Animal Care and Use Committee.

#### Consent for publication

Authors consent for the publication. All authors read and approved the final manuscript.

#### Competing interests

The authors declare no conflicts of interest.

Received: 30 June 2023 / Accepted: 20 May 2024

Published online: 09 June 2024

### References

1. 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, Kretschmar HA, Trojanowski JQ, Lee VM-Y. Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and amyotrophic lateral sclerosis. *Science*. 2006;314:130–3.
2. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun*. 2006;351:602–11.
  3. Hayes LR, Kalab P. Emerging therapies and novel targets for TDP-43 proteinopathy in ALS/FTD. *Neurotherapeutics*. 2022;19:1061–84.
  4. Meneses A, Koga S, O'Leary J, Dickson DW, Bu G, Zhao N. TDP-43 Pathology in Alzheimer's Disease. *Mol Neurodegener*. 2021;16:84.
  5. Nelson PT, Dickson DW, Trojanowski JQ, Jack CR, Boyle PA, Arfanakis K, Rademakers R, Alafuzoff I, Attems J, Brayne C, Coyle-Gilchrist ITS, Chui HC, Fardo DW, Flanagan ME, Halliday G, Hokkanen SRK, Hunter S, Jicha GA, Katsumata Y, Kawas CH, Keene CD, Kovacs GG, Kukull WA, Levey AI, Makkinejad N, Montine TJ, Murayama S, Murray ME, Nag S, Rissman RA, Seeley WW, Sperling RA, Schneider CLW. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain*. 2019;142:1503–1527.
  6. Watts GDJ, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP, Kimonis VE. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet*. 2004;36:377–81.
  7. Britton KA, Ling JP, Braunstein KE, Montagne JM, Kastenschmidt JM, Wilson A, Ikenaga C, Tsao W, Pinal-Fernandez I, Russell KA, Reed N, Mozaffar T, Wagner KR, Ostrow LW, Corse AM, Mammen AL, Villalta SA, Larman HB, Wong PC, Lloyd TE. Loss of TDP-43 function and rimmed vacuoles persist after T cell depletion in a xenograft model of sporadic inclusion body myositis. *Sci Transl Med*. 2022;14:eabi9196.
  8. Neumann M, Lee EB, Mackenzie IR. Frontotemporal dementias, emerging milestones of the 21st Century. *Adv Exp Med Biol*. 2021;1281:201–17.
  9. Sephton CF, Good SK, Atkin S, Dewey CM, Mayer P, Herz J, Yu G. TDP-43 is a developmentally regulated protein essential for early embryonic development. *J Biol Chem*. 2010;285:38740.
  10. Kraemer BC, Schuck T, Wheeler JM, Robinson LC, Trojanowski JQ, Lee VMY, Schellenberg GD. Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis. *Acta Neuropathol*. 2010;119:409–19.
  11. Chiang P-M, Ling J, Jeong YH, Price DL, Aja SM, Wong PC. Deletion of TDP-43 down-regulates Tbc1d11, a gene linked to obesity, and alters body fat metabolism. *Proc Natl Acad Sci*. 2010;107:16320–4.
  12. Yang C, Wang H, Qiao T, Yang B, Aliaga L, Qiu L, Tan W, Salameh J, McKenna-Yasek DM, Smith T, Peng L, Moore MJ, Brown RH, Cai H, Xu Z. Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proc Natl Acad Sci*. 2014;111:E1121–9.
  13. Feiguin F, Godena VK, Romano G, D'Ambrogio A, Klima R. Baralle, depletion of TDP-43 affects *Drosophila* motoneurons terminal synapsis and locomotive behavior. *FEBS Lett*. 2009;583:1586–92.
  14. Schmid B, Hruscha A, Hogl S, Banzhaf-Strathmann J, Strecker K, van der Zee J, Teucke M, Eimer S, Hegermann J, Kittelmann M, Kremmer E, Cruts M, Solchenberger B, Hasenkamp L, van Bebber F, Broeckhoven CV, Edbauer D, Lichtenthaler SF, Haass C. Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth. *Proc Natl Acad Sci*. 2013;110:4986–91.
  15. Ling JP, Pletnikova O, Troncoso JC, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 2015;349:650–5.
  16. Jeong YH, Ling JP, Lin SZ, Donde AN, Braunstein KE, Majounie E, Traynor BJ, LaClair KD, Lloyd TE, Wong PC. Tdp-43 cryptic exons are highly variable between cell types. *Mol Neurodegener*. 2017;12:13.
  17. Klim JR, Williams LA, Limone F, Juan JGS, 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. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci*. 2019;22:167–79.
  18. Melamed Z, López-Erauskin J, Baughn MW, Zhang O, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, Ohkubo T, Rodriguez M, Lin N, Wu D, Bennett CF, Rigo F, Cruz SD, Ravits J, Lagier-Tourenne C, Cleveland DW. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*. 2019;22:180–90.
  19. Baughn MW, Melamed Z, López-Erauskin J, Beccari MS, Ling K, Zuberi A, Presa M, Gonzalo-Gil E, Maimon R, Vazquez-Sanchez S, Chaturvedi S, Bravo-Hernández M, Taupin V, Moore S, Artates JW, Acks E, Ndayambaje IS, de Quadros AR, Jafar-nejad P, Rigo F, Bennett CF, Lutz C, Lagier-Tourenne C, Cleveland DW. Mechanism of STMN2 cryptic splice-polyadenylation and its correction for TDP-43 proteinopathies. *Science*. 2023;379:1140–9.
  20. Prudencio M, Humphrey J, Pickles S, Brown A-L, Hill SE, Kachergus JM, Shi J, Heckman MG, Spiegel MR, Cook C, Song Y, Yue M, Daugherty LM, Carlomagno Y, Jansen-West K, de Castro CF, DeTure M, Koga S, Wang Y-C, Sivakumar P, Bodo C, Candalija A, Talbot K, Selvaraj BT, Burr K, Chandran S, Newcombe J, Lashley T, Hubbard I, Catalano D, Kim D, Propp N, Fennessey S, Consortium NA, Fagegaltier D, Phatnani H, Secrier M, Fisher EM, Oskarsson B, van Blitterswijk M, Rademakers R, Graff-Radford NR, Boeve BF, Knopman DS, Petersen RC, Josephs KA, Thompson EA, Raj T, Ward M, Dickson DW, Gendron TF, Fratta P, Petrucelli L. Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia. *J Clin Investig*. 2020;130:6080–92.
  21. Koike Y, Pickles S, Ayuso VE, Jansen-West K, Qi YA, Li Z, Daugherty LM, Yue M, Zhang Y-J, Cook CN, Dickson DW, Ward M, Petrucelli L, Prudencio M. TDP-43 and other hnRNPs regulate cryptic exon inclusion of a key ALS/FTD risk gene, UNC13A. *PLOS Biol*. 2023;21:e3002028.
  22. Brown A-L, Wilkins OG, Keuss MJ, Hill SE, Zanovello M, Lee WC, Bampton A, Lee FCY, Masino L, Qi YA, Bryce-Smith S, Gatt A, Hallegger M, Fagegaltier D, Phatnani H, Phatnani H, Kwan J, Sareen D, Broach JR, Simmons Z, Arcila-Londono X, Lee EB, Deerlin VMV, Shneider NA, Fraenkel E, Ostrow LW, Baas F, Zaitlen N, Berry JD, Malaspina A, Fratta P, Cox GA, Thompson LM, Finkbeiner S, Dardiotis E, Miller TM, Chandran S, Pal S, Hornstein E, MacGowan DJ, Heiman-Patterson T, Hammell MG, Patsopoulos NA, Butovsky O, Dubnau J, Nath A, Bowser R, Harms M, Aronica E, Poss M, Phillips-Cremens J, Cray J, Atassi N, Lange DJ, Adams DJ, Stefanis L, Gotkine M, Baloh RH, Babu S, Raj T, Paganoni S, Shalem O, Smith C, Zhang B, Harris B, Broce I, Drory V, Ravits J, McMillan C, Menon V, Wu L, Altschuler S, Lerner Y, Sattler R, Keuren-Jensen KV, Rozenblatt-Rosen O, Lindblad-Toh K, Nicholson K, Gregersen P, Lee J-H, Kokos S, Muljo S, Newcombe J, Gustavsson EK, Seddighi S, Reyes JF, Coon SL, Ramos D, Schiavo G, Fisher EMC, Raj T, Secrier M, Lashley T, Ule J, Buratti E, J. Humphrey, M. E. Ward, P. Fratta, TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. *Nature*. 2022;603:1317137.
  23. Ma XR, Prudencio M, Koike Y, Vatsavayai SC, Kim G, Harbinski F, Briner A, Rodriguez CM, Guo C, Akiyama T, Schmidt HB, Cummings BB, Wyatt DW, Kurylo K, Miller G, Mekhoubad S, Sallee N, Mekonnen G, Ganser L, Rubien JD, Jansen-West K, Cook CN, Pickles S, Oskarsson B, Graff-Radford NR, Boeve BF, Knopman DS, Petersen RC, Dickson DW, Shorter J, Myong S, Green EM, Seeley WW, Petrucelli L, Gitler AD. TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A. *Nature*. 2022;603:124–30.
  24. Sun M, Bell W, LaClair KD, Ling JP, Han H, Kageyama Y, Pletnikova O, Troncoso JC, Wong PC, Chen LL. Cryptic exon incorporation occurs in Alzheimer's brain lacking TDP-43 inclusion but exhibiting nuclear clearance of TDP-43. *Acta Neuropathol*. 2017;133:923–31.
  25. Tan Q, Yalamanchili HK, Park J, Maio AD, Lu H-C, Wan Y-W, White JJ, Bondar VV, Sayegh LS, Liu X, Gao Y, Sillitoe RV, Orr HT, Liu Z, Zoghbi HY. Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum Mol Genet*. 2016;25:ddw337.
  26. Torres P, Ramirez-Núñez O, Romero-Guevara R, Barés G, Granado-Serrano AB, Ayala V, Boada J, Fontdevila L, Povedano M, Sanchis D, Pamplona R, Ferrer I, Portero-Otín, cryptic exon splicing function of TARDBP interacts with autophagy in nervous tissue. *Autophagy*. 2018;14:1398–403.
  27. Roczniak-Ferguson A, Ferguson SM. Pleiotropic requirements for human TDP-43 in the regulation of cell and organelle homeostasis. *Life Sci Alliance*. 2019;2:e201900358.
  28. Šušnjarić U, Škrabar N, Brown A-L, Abbasi Y, Phatnani H, Consortium NA, Phatnani H, Fratta P, Kwan J, Sareen D, Broach JR, Simmons Z, Arcila-Londono X, Lee EB, Deerlin VMV, Shneider NA, Fraenkel E, Ostrow LW, Baas F, Berry JD, Butovsky O, Baloh RH, Shalem O, Heiman-Patterson T, Stefanis L, Chandran S, Pal S, Smith C, Malaspina A, Hammell MG, Patsopoulos NA, Dubnau J, Poss M, Zhang B, Zaitlen N, Hornstein E, Miller TM, Dardiotis E, Bowser R, Menon V, Harms M, Atassi N, Lange DJ, MacGowan DJ, McMillan C, Aronica E, Harris B, Ravits J, Cray J, Thompson LM, Raj T, Paganoni S, Adams DJ, Babu S, Drory V, Gotkine M, Broce I, Phillips-Cremens J, Nath A, Finkbeiner S, Cox GA, Cortese A, Cereda C, Bugiardini E, Cardani R, Meola G, Ripolone M, Moggio M, Romano M, Secrier M, Fratta P. E. Buratti, Cell environment shapes TDP-43 function with implications in neuronal and muscle disease. *Commun Biol*. 2022;5:314.
  29. Mehta PR, Brown A-L, Ward ME, Fratta P. The era of cryptic exons: implications for ALS-FTD. *Mol Neurodegener*. 2023;18:16.
  30. Tsao W, Jeong YH, Lin S, Ling J, Price DL, Chiang P-M, Wong PC. Rodent models of TDP-43: recent advances. *Brain Res*. 2012;1462:26–39.

31. Shan X, Chiang P-M, Price DL, Wong PC. Altered distributions of Gemini of coiled bodies and mitochondria in motor neurons of TDP-43 transgenic mice. *Proc Natl Acad Sci*. 2010;107:16325716330.
32. Wils H, Kleinberger G, Janssens J, Pereson S, Joris G, Cuijt I, Smits V, Groote CC, Broeckhoven CV, Kumar-Singh S. TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration. *Proc Natl Acad Sci*. 010;107:385873863.
33. Stallings NR, Puttappathi K, Luther CM, Burns DK, Elliott JL. Progressive motor weakness in transgenic mice expressing human TDP-43. *Neurobiol Dis*. 2010;40:404–14.
34. Wegorzewska I, Bell S, Cairns NJ, Miller TM, Baloh RH. TDP-43 mutant transgenic mice develop features of ALS and frontotemporal lobar degeneration. *Proc Natl Acad Sci*. 2009;106:18809–18814.
35. Xu Y-F, Gendron TF, Zhang Y-J, Lin W-L, D'Alton S, Sheng H, Casey MC, Tong J, Knight J, Yu X, Rademakers R, Boylan K, Hutton M, McGowan E, Dickson DW, Lewis J, Petrucelli L. Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J Neurosci: off J Soc Neurosci*. 2010;30:10851–9.
36. Xu Y-F, Zhang Y-J, Lin W-L, Cao X, Stetler C, Dickson DW, Lewis J, Petrucelli L. Expression of mutant TDP-43 induces neuronal dysfunction in transgenic mice. *Mol Neurodegener*. 2011;6:73–73.
37. Tsai K-J, Yang C-H, Fang Y-H, Cho K-H, Chien W-L, Wang W-T, Wu T-W, Lin C-P, Fu W-M, Shen C-KJ. Elevated expression of TDP-43 in the forebrain of mice is sufficient to cause neurological and pathological phenotypes mimicking FTL-D-U. *J Exp Med*. 2010;207:1661–73.
38. Igaz LM, Kwong LK, Chen-Plotkin A, Winton MJ, Unger TL, Xu Y, Neumann M, Trojanowski JQ, Lee VM-Y. Expression of TDP-43 C-terminal fragments in vitro recapitulates pathological features of TDP-43 Proteinopathies\*. *J Biol Chem*. 2009;284:8516–24.
39. Swarup V, Phaneuf D, Bareil C, Robertson J, Rouleau GA, Kriz J, Julien J-P. Pathological hallmarks of amyotrophic lateral sclerosis/frontotemporal lobar degeneration in transgenic mice produced with TDP-43 genomic fragments. *Brain*. 2011;134:2610–26.
40. Zhou H, Huang C, Chen H, Wang D, Landel CP, Xia PY, Bowser R, Liu Y-J, Xia XG. Transgenic rat model of Neurodegeneration caused by mutation in the TDP Gene. *PLoS Genet*. 2010;6:e1000887.
41. Janssens J, Wils H, Kleinberger G, Joris G, Cuijt I, Groote CC, Broeckhoven CV, Kumar-Singh S. Overexpression of ALS-Associated p.M337V Human TDP-43 in mice worsens Disease features compared to wild-type human TDP-43 mice. *Mol Neurobiol*. 2013;48:22–35.
42. Sieverding K, Ulmer J, Bruno C, Satoh T, Tsao W, Freischmidt A, Akira S, Wong PC, Ludolph AC, Danzer KM, Lobsiger CS, Brenner D, Weishaupt JH. Hemizygous deletion of Tdk1 worsens neuromuscular junction pathology in TDP-43 transgenic mice. *Exp Neurol*. 2021;335:113496.
43. Hergesheimer RC, Chami AA, de Assis DR, Your'c'h P, Andres CR, Corcia P, Lanzaster D, Blasco H. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain*. 2019;142:1176–94.
44. Yang C, Qiao T, Yu J, Wang H, Guo Y, Salameh J, Metterville J, Parsi S, Yusuf I, Brown RH, Cai H, Xu Z. Low-level overexpression of wild type TDP-43 causes late-onset, progressive neurodegeneration and paralysis in mice. *PLoS ONE*. 2022;17:e0255710.
45. Philips T, Rothstein JD. Rodent Models of Amyotrophic Lateral Sclerosis. *Curr Protoc Pharmacol*. 2015;69:5.67.1-5.67.21.
46. Ayala YM, Conti LD, Avendaño-Vázquez SE, Dhir A, Romano M, D'Ambrogio A, Tollervey J, Ule J, Baralle M, Buratti E, Baralle FE. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J*. 2011;30:277–88.
47. Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling S-C, Sun E, Wancewicz E, Mazur C, Kordasiewicz H, Sedaghat Y, Donohue JP, Shiu L, Bennett CF, Yeo GW, Cleveland DW. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14:459–68.
48. Avendaño-Vázquez SE, Dhir A, Bembich S, Buratti E, Proudfoot N, Baralle FE. Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev*. 2012;26:1679–84.
49. Xie M, Liu YU, Zhao S, Zhang L, Bosco DB, Pang Y-P, Zhong J, Sheth U, Martens YA, Zhao N, Liu C-C, Zhuang Y, Wang L, Dickson DW, Mattson MP, Bu G, Wu L-J. TREM2 interacts with TDP-43 and mediates microglial neuroprotection against TDP-43-related neurodegeneration. *Nat Neurosci*. 2022;25:26–38.
50. Spiller KJ, Restrepo CR, Khan T, Dominique MA, Fang TC, Canter RG, Roberts C, Miller KR, Ransohoff RM, Trojanowski JQ, Lee VM-Y. Microglia-mediated recovery from ALS-relevant motor neuron degeneration in a mouse model of TDP-43 proteinopathy. *Nat Neurosci*. 2018;21:329–40.
51. Hunter M, Spiller KJ, Dominique MA, Xu H, Hunter FW, Fang TC, Canter RG, Roberts CJ, Ransohoff RM, Trojanowski JQ, Lee VM-Y. Microglial transcriptome analysis in the rNLS8 mouse model of TDP-43 proteinopathy reveals discrete expression profiles associated with neurodegenerative progression and recovery. *Acta Neuropathol Commun*. 2021;9:140.
52. Porta S, Xu Y, Restrepo CR, Kwong LK, Zhang B, Brown HJ, Lee EB, Trojanowski JQ, Lee VM-Y. Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat Commun*. 2018;9:4220.
53. Spiller KJ, Cheung CJ, Restrepo CR, Kwong LK, Stieber AM, Trojanowski JQ, Lee VM-Y. Selective Motor Neuron Resistance and Recovery in a New Inducible Mouse Model of TDP-43 Proteinopathy. *J Neurosci*. 2016;36:7707–17.
54. Igaz LM, Kwong LK, Lee EB, Chen-Plotkin A, Swanson E, Unger T, Malunda J, Xu Y, Winton MJ, Trojanowski JQ, Lee VM-Y. Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. *J Clin Invest*. 2011;121:726–38.
55. Arnold ES, Ling S-C, Huelga SC, Lagier-Tourenne C, Polymenidou M, Ditsworth D, Kordasiewicz HB, McAlonis-Downes M, Platoshyn O, Parone PA, Cruz SD, Clutario KM, Swing D, Tessarollo L, Marsala M, Shaw CE, Yeo GW. D. W. Cleveland, ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proc Natl Acad Sci*. 2013;110:E7367E745.
56. Fratta P, Sivakumar P, Humphrey J, Lo K, Ricketts T, Oliveira H, Brito-Armas JM, Kalmar B, Ule A, Yu Y, Birsa N, Bodo C, Collins T, Conicella AE, Maza AM, Marrero-Gagliardi A, Stewart M, Mianne J, Corrochano S, Emmett W, Codner G, Groves M, Fukumura R, Gondo Y, Lythgoe M, Pauws E, Peskett E, Stanier P, Teboul L, Hallegger M, Calvo A, Chiò A, Isaacs AM, Fawzi NL, Wang E, Housman DE, Baralle F, Greensmith L, Buratti E, Plagnol V, Fisher EM, Acevedo-Arozena A. Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J*. 2018;37:e98684.
57. White MA, Kim E, Duffy A, Adalbert R, Phillips BU, Peters OM, Stephenson J, Yang S, Massenzio F, Lin Z, Andrews S, Segonds-Pichon A, Metterville J, Saksida LM, Mead R, Ribchester RR, Barhomi Y, Serre T, Coleman MP, Fallon JR, Bussey TJ, Brown RH, Sreedharan J. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat Neurosci*. 2018;21:552–63.
58. Walker AK, Spiller KJ, Ge G, Zheng A, Xu Y, Zhou M, Tripathy K, Kwong LK, Trojanowski JQ, Lee VM-Y. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. *Acta Neuropathol*. 2015;130:643–60.
59. McClory SP, Lynch KW, Ling JP. HnRNP L represses cryptic exons. *RNA*. 2018;24:761–8.
60. Ebstein SY, Yagudayeva I, Shneider NA. Mutant TDP-43 causes early-stage dose-Dependent Motor Neuron Degeneration in a TARDBP Knockin Mouse Model of ALS. *Cell Rep*. 2019;26:364–e3734.
61. Gordon D, Dafinca R, Scaber J, Alegre-Abarategui J, Farrimond L, Scott C, Biggs D, Kent L, Oliver PL, Davies B, Ansong O, Wade-Martins R, Talbot K. Single-copy expression of an amyotrophic lateral sclerosis-linked TDP-43 mutation (M337V) in BAC transgenic mice leads to altered stress granule dynamics and progressive motor dysfunction. *Neurobiol Dis*. 2019;121:148–62.
62. Stribl C, Samara A, Trümbach D, Peis R, Neumann M, Fuchs H, Gailus-Durner V, de Angelis MH, Rathkolb B, Wolf E, Beckers J, Horsch M, Neff F, Kremmer E, Koob S, Reichert AS, Hans W, Rozman J, Klingenspor M, Aichler M, Walch AK, Becker L, Klopstock T, Glasl L, Hölter SM, Wurst W, Floss T. Mitochondrial dysfunction and decrease in Body Weight of a Transgenic knock-in mouse model for TDP-43\*. *J Biol Chem*. 2014;289:10769–84.
63. Pesiridis GS, Lee VM-Y, Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009;18:R156–62.
64. Fernandopulle MS, Prestil R, Grunseich C, Wang C, Gan L, Ward ME. Transcription factor-mediated differentiation of human iPSCs into neurons. *Curr Protoc Cell Biol*. 2018;79:e51.
65. ENCODE Project Consortium et al. "Expanded encyclopaedias of DNA elements in the human and mouse genomes." *Nature* vol. 2020;583:7818:699–710. <https://doi.org/10.1038/s41586-020-2493-4>.
66. Krus KL, Strickland A, Yamada Y, Devault L, Schmidt RE, Bloom AJ, Milbrandt J, DiAntonio A. Loss of Stathmin-2, a hallmark of TDP-43-associated ALS, causes motor neuropathy. *Cell Rep*. 2022;39:111001.



67. Chung M, Carter EK, Veire AM, Dammer EB, Chang J, Duong DM, Raj N, Bas-sell GJ, Glass JD, Gendron TF, Nelson PT, Levey AI, Seyfried NT, McEachin ZT. Cryptic exon inclusion is a molecular signature of LATE-NC in aging brains. *Acta Neuropathol.* 2024;147:29.
68. Chang K, Ling JP, Redding-Ochoa J, An Y, Li L, Dean SA, Blanchard TG, Pylyukh T, Barrett A, Irwin KE, Moghekar A, Resnick SM, Wong PC. Troncoso, loss of TDP-43 splicing repression occurs early in the aging population and is associated with Alzheimer's disease neuropathologic changes and cognitive decline. *Acta Neuropathol.* 2024;147:4.
69. Wilks C, Gaddipati P, Nellore A, Langmead B. Snaptron: querying splicing patterns across tens of thousands of RNA-seq samples. *Bioinformatics.* 2018;34:114–6.
70. Wilks C, Zheng SC, Chen FY, Charles R, Solomon B, Ling JP, Imada EL, Zhang D, Joseph L, Leek JT, Jaffe AE, Nellore A, Collado-Torres L, Hansen KD, Langmead B. recount3: summaries and queries for large-scale RNA-seq expression and splicing. *Genome Biol.* 2021;22:323.
71. Ling JP, Wilks C, Charles R, Leavey PJ, Ghosh D, Jiang L, Santiago CP, Pang B, Venkataraman A, Clark BS, Nellore A, Langmead B, Blackshaw S. ASCOT identifies key regulators of neuronal subtype-specific splicing. *Nat Commun.* 2020;11:137.
72. Ling JP, Bygrave AM, Santiago CP, Carmen-Orozco RP, Trinh VT, Yu M, Li Y, Liu Y, Bowden KD, Duncan LH, Han J, Taneja K, Dongmo R, Babola TA, Parker P, Jiang L, Leavey PJ, Smith JJ, Vistein R, Gimmen MY, Dubner B, Helmenstine E, Teodorescu P, Karantanos T, Ghiaur G, Kanold PO, Bergles D, Langmead B, Sun S, Nielsen KJ, Peachey N, Singh MS, Dalton WB, Rajaii F, Huganir RL, Blackshaw S. Cell-specific regulation of gene expression using splicing-dependent frameshifting. *Nat Commun.* 2022;13:5773.
73. Gte T, Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369:1318–30.
74. Hans F, Fiesel FC, Strong JC, Jäckel S, Rasse TM, Geisler S, Springer W, Schulz JB, Voigt A, Kahle PJ. UBE2E ubiquitin-conjugating enzymes and Ubiquitin Isopeptidase Y regulate TDP-43 protein Ubiquitination\*. *J Biol Chem.* 2014;289:19164–79.
75. Cohen TJ, Hwang AW, Restrepo CR, Yuan C-X, Trojanowski JQ, Lee VMY. An acetylation switch controls TDP-43 function and aggregation propensity. *Nat Commun.* 2015;6:5845.
76. Wang P, Wander CM, Yuan C-X, Bereman MS, Cohen TJ. Acetylation-induced TDP-43 pathology is suppressed by an HSF1-dependent chaperone program. *Nat Commun.* 2017;8:82.
77. Hans F, Eckert M, von Zweydoorf F, Gloeckner CJ, Kahle PJ. Identification and characterization of ubiquitylation sites in TAR DNA-binding protein of 43 kDa (TDP-43). *J Biol Chem.* 2018;293:16083–99.
78. Mihevc SP, Baralle M, Buratti E, Rogelj B. TDP-43 aggregation mirrors TDP-43 knockdown, affecting the expression levels of a common set of proteins. *Sci Rep.* 2016;6:33996.
79. Mann JR, Gleixner AM, Mauna JC, Gomes E, DeChellis-Marks MR, Needham PG, Copley KE, Hurtle B, Portz B, Pyles NJ, Guo L, Calder CB, Wills ZP, Pandey UB, Kofler JK, Brodsky JL, Thathiah A, Shorter J, Donnelly, RNA binding antagonizes neurotoxic phase transitions of TDP-43. *Neuron.* 2019;102:321–e3388.
80. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohli SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstern S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596:583–9.
81. Duan L, Zaepfel BL, Aksenova V, Dasso M, Rothstein JD, Kalab P, Hayes LR. Nuclear RNA binding regulates TDP-43 nuclear localization and passive nuclear export. *Cell Rep.* 2022;40:111106–111106.
82. 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).
83. Chen H-J, Topp SD, Hui HS, Zacco E, Katarya M, McLoughlin C, King A, Smith BN, Troakes C, Pastore A, Shaw CE. RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. *Brain.* 2019;142:3753–70.
84. Agrawal S, Jain M, Yang W, Yuan HS. Frontotemporal dementia-linked P112H mutation of TDP-43 induces protein structural change and impairs its RNA binding function. *Protein Sci.* 2021;30:350–65.
85. Hruska-Plochan M, Wiersma VI, Betz KM, Mallona I, Ronchi S, Maniecka Z, Hock E-M, Tantardini E, Laferriere F, Sahadevan S, Hoop V, Delvendahl I, Pérez-Berlanga M, Gatta B, Panatta M, van der Bourg A, Bohaciakova D, Sharma P, Vos LD, Frontzek K, Aguzzi A, Lashley T, Robinson MD, Karayannis T, Mueller M, Hierlemann A, Polymenidou M. A model of human neural networks reveals NPTX2 pathology in ALS and FTLD. *Nature.* 1–11 (2024).
86. Wang C, Ward ME, Chen R, Liu K, Tracy TE, Chen X, Xie M, Sohn PD, Ludwig C, Meyer-Franke A, Karch CM, Ding S, Gan L. Scalable production of iPSC-Derived human neurons to identify tau-lowering compounds by high-content screening. *Stem Cell Rep.* 2017;9:1221–33.
87. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15–21.
88. Wilks C, Ahmed O, Baker DN, Zhang D, Collado-Torres L, Langmead B. Mega-depth: efficient coverage quantification for BigWigs and BAMs. *Bioinformatics.* 2021;37:3014–6.
89. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The Human Genome Browser at UCSC. *Genome Res.* 2002;12:996–1006.
90. Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, Sisu C, Wright JC, Armstrong J, Barnes I, Berry A, Bignell A, Boix C, Carbonell Sala S, Cunningham F, Di Domenico T, Donaldson S, Fiddes IT, García Girón C, Gonzalez JM, Grego T, Hardy M, Hourlier T, Howe KL, Hunt T, Izuogu OG, Johnson R, Martin FJ, Martínez L, Mohanan S, Muir P, Navarro FCP, Parker A, Pei B, Pozo F, Riera FC, Ruffier M, Schmitt BM, Stapleton E, Suner M-M, Sycheva I, Uszczynska-Ratajczak B, Wolf MY, Xu J, Yang YT, Yates A, Zerbino D, Zhang Y, Choudhary JS, Gerstein M, Guigó R, T. J. P., Hubbard M, Kellis B, Paten P, Flicek. GENCODE 2021. *Nucleic Acids Res.* 2020;49:gkaa1087-.
91. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Židek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M, Cowie A, Hobbs N, Kohli P, Kleywegt G, Birney E, Hassabis D, Velankar S. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2021;50:D439–44.

## Publisher's Note

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