Drivers of TDP43 Dyshomeostasis in Amyotrophic Lateral Sclerosis

by

Kaitlin T. Weskamp

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Neuroscience)
in the University of Michigan
2019

Doctoral Committee:

Assistant Professor Sami Barmada, Chair
Professor Geoffrey Murphy
Professor Jack Parent
Associate Professor Peter Todd
Professor Michael Uhler

Kaitlin T. Weskamp

kweskamp@umich.edu

ORCID iD: 0000-0003-4474-2850

Dedication

This dissertation is dedicated to John, who has offered unwavering love, support, and confidence in my abilities through every academic endeavor.

Acknowledgements

First, I thank my mentor Dr. Sami Barmada, for his expertise, guidance, and patience throughout my training. He has done a remarkable job of curating a safe and supportive environment in which his students can grow as scientists, and I am thankful for the opportunity to work with him. I thank members of the Barmada laboratory for their contributions to our hilarious, bizarre, and productive work environment, and for making it a joy to come to work every day. These relationships are among the most valuable things I gained during my time in graduate school, and I consider myself lucky to work among friends. I thank my support system, particularly John, Anita, and Megan, who kept me moving through the lows and reminded me to celebrate the highs. I thank my dissertation committee members Dr. Michael Uhler, Dr. Jack Parent, Dr. Geoffrey Murphy, and Dr. Peter Todd for their valuable suggestions that strengthened the quality of my work. I thank members of the Feldman, Paulson, and Todd labs for their generosity with expertise, reagents, and hallway conversations. I thank the staff of the Neuroscience Graduate Program, the Program in Biomedical Sciences, the Neurology Department, and the Biomedical Sciences Research Building janitorial staff for all their hard work over the years. Other collaborators, contributors, and sources of support are acknowledged throughout the dissertation

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disorder in which the progressive loss of motor neurons results in paralysis and respiratory failure. Though the study of ALS is complicated by its heterogeneous biochemical, genetic, and clinical features, dysregulation of the RNA-binding protein TDP43 is observed in the vast majority of ALS cases. Although *TARDBP* mutations account for only a small proportion of the disease burden (2-5%), cytoplasmic TDP43 mislocalization and accumulation are observed in >90% of individuals with ALS. Moreover, mutations in several other ALS-associated genes result in TDP43 pathology. TDP43 is an essential protein involved in several RNA processing events, including splicing, translation, and degradation, and small changes in its localization and expression level are sufficient to disrupt critical cell processes (Chapter 1). As such, accumulating evidence implicates TDP43 and TDP43-dependent RNA processing in neurodegenerative disease (Chapter 2), but drivers of TDP43 accumulation and mislocalization remain fundamentally unclear. Here, we seek to identify phenomena that initiate TDP43 dyshomeostasis and develop techniques to better monitor TDP43 metabolism in the context of ALS.

Much like TDP43 pathology, neuronal hyperexcitability is a conserved feature observed in both familial and sporadic ALS. However, its relation to neurodegeneration and TDP43 deposition in disease remains unknown. In Chapter 3, we show that hyperexcitability recapitulates TDP43 pathology by upregulating shortened (s) TDP43 splice isoforms. These truncated isoforms accumulate in the cytosol, where they form insoluble inclusions that sequester full-length TDP43

via preserved N-terminal interactions. Consistent with these findings, sTDP43 overexpression is toxic to mammalian neurons, suggesting that neurodegeneration results from complementary gain- and loss-of-function mechanisms. In humans and mice, sTDP43 transcripts are enriched in vulnerable motor neurons, and we observed a striking accumulation of sTDP43 within neurons and glia of ALS patients. These studies uncover a hitherto unknown role of alternative TDP43 isoforms, and indicate that sTDP43 production may be a key contributor to the susceptibility of motor neurons in ALS.

In Appendix A, we establish a technique to monitor TDP43 metabolism at the endogenous level. To do so, we developed induced pluripotent stem cell (iPSC)-derived neurons in which we can monitor the synthesis and degradation of native TDP43 in a non-invasive manner. Following these measurements, each neuron is tracked over time to determine its time of death via longitudinal fluorescence microscopy (Appendix B), enabling us to determine how TDP43 synthesis and decay rates impact neuronal survival. Future work can utilize this methodology to determine if TDP43 metabolism is altered in neurons derived from ALS patients with *C9orf72* and *TARDBP* mutations to further elucidate mechanisms of TDP43 dyshomeostasis.

Chapter 4 concludes the dissertation and describes future studies to better understand mechanisms of sTDP43 toxicity and determine if sTDP43 is a viable therapeutic target for ALS. Appendix C further explores the identification of novel therapies and the development of a medium-throughput screen to identify novel compounds that stop or attenuate neurodegeneration. Taken together, this dissertation uncovers a novel disease pathway that may be targeted for therapeutic development, and establishes a technique to determine how TDP43 dyshomeostasis contributes to neurodegeneration in ALS.

Chapter 1. Introduction*

1.1 Overview

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder in which the progressive loss of motor neurons results in paralysis and respiratory failure[1]. The pattern and rate of symptom progression varies widely between patients, but ALS is typically fatal within 3-5 years of onset. While 90-95% of ALS cases are sporadic and have no known genetic cause, the remaining 5-10% are associated with heritable mutations in one of over 15 different genes; a list that grows steadily longer with technological advances in whole genome sequencing and an increasing trend of pooling data between international consortia[2,3].

Despite significant effort, there is currently no effective disease-modifying therapy for ALS. Over 60 molecules have been investigated as potential therapeutics, but the vast majority of human trials have failed to achieve clinical efficacy[4]. Only two drugs are approved for the treatment of ALS: the first is Riluzole, an anti-glutamatergic agent that extends patient lifespan by an average of 2-3 months[5], and the second is the anti-oxidative treatment Edaravone, which slows disease progression in patients in the earlier stages of disease[6,7]. However, these drugs do not alter disease progression or significantly extend the lifespan of ALS patients. This marked lack of progress is potentially due to the clinical and genetic heterogeneity associated with ALS, and as

* A portion of this chapter represents the following manuscript:

Weskamp K, Barmada SJ. (2018). TDP43 and RNA instability in amyotrophic lateral sclerosis. *Brain Research.* https://doi.org/10.1016/j.brainres.2018.01.015

such we aim to identify and understand convergent disease mechanisms common to the majority of patients with this disorder.

It is increasingly clear that RNA dysregulation is a key contributor to ALS pathogenesis. Over the past decade, disease-associated mutations have been identified in genes encoding multiple RNA-binding proteins participating in all aspects of RNA processing[8]. Among these is TDP43, a nuclear protein integrally involved in RNA metabolism. Although mutations in the gene encoding TDP43 (*TARDBP*) account for only a small proportion of the disease burden (2-5%), cytoplasmic TDP43 mislocalization and accumulation are observed in >90% of individuals with ALS[9]. Moreover, mutations in several other ALS-associated genes — including *C9orf72*[10], *ANG*[11], *TBK1*[12], *PFN1*[13], *UBQLN2*[14], *VCP*[15], and *hnRNPA2/B1*[16] — result in TDP43 pathology. This convergence heavily implicates TDP43 and TDP43-dependent RNA processing in ALS pathogenesis, and suggests that a better understanding of how it contributes to neurodegeneration may reveal targets for more effective therapies.

1.2 TDP43 Pathology in ALS

Although TDP43 is a primarily nuclear protein, its clearance from the nucleus and accumulation in cytosolic aggregates is observed in >90% of ALS cases[17,18]. These cytosolic inclusions generally appear as either filamentous skeins or compact inclusion bodies[19,20] within motor neurons and glia in the brain and spinal cord[17], and there is some evidence that the spread of TDP43 pathology across cortical regions can be used to track ALS progression[21]. Extensive work in patient tissue shows that these inclusions are insoluble, ubiquitinated[17], and hyperphosphorylated[22,23], suggesting that TDP43 undergoes a histopathological transformation in disease. Moreover, phosphorylated C-terminal fragments of TDP43 are present within these

inclusions in the brain, and to a lesser extent in the spinal cord[24]. Despite thorough characterization of this pathology, how or if the mislocalization and accumulation of TDP43 drives ALS pathogenesis remains unclear.

1.3 TDP43 Structure

TDP43 is encoded by the gene *TARDBP*, and was initially identified as a transcriptional repressor of HIV[25] prior to its association with ALS in 2006[17,26]. The protein itself consists of 414 amino acids that encode several functional domains (Figure 1.1). The N-terminus (amino acids (aa) 1-102) is involved in TDP43 self-association, and regulates the formation of TDP43 homodimers[27,28] that may be important for normal protein function[28]. It also contains a nuclear localization signal (NLS, aa 82-98) that regulates trafficking between the nucleus and the cytoplasm[29,30]. TDP43 also includes two highly-conserved RNA-recognition motifs, RRM1 (aa 104-176) and RRM2 (aa 192-262), that widely bind both DNA and RNA with higher specificity towards UG/TG-rich sequences[31–33]. These RRMs regulate several RNA metabolic processes, such as mRNA processing, export, and stability[31,34,35]. The C-terminal region (aa 277-414) encompasses a prion-like domain (aa 345-366) and a glycine-rich region (aa 366-414), and serves as a critical regulator of protein-protein interactions[36,37], alternative splicing[34,37,38], and localization[30]. This region is intrinsically disordered and canonically aggregation-prone[39], though emerging evidence suggests that it is a key regulator of protein solubility and folding in healthy cells[40].

1.4 Other TDP43 Isoforms

The *TARDBP* gene is comprised of six exons, and several groups have described alternative splicing events across the transcript. To date, nineteen distinct transcripts have been reported[41,42].

The majority arise from splicing events within exon 6 and the 3' UTR, with some leading to the inclusion of a previously unannotated exon 7[41–44]. Though the majority of these isoforms are not well characterized, some work has shown that a subset form insoluble, ubiquitinated, cytosolic inclusions and recapitulate key features of TDP43 pathology in ALS[45,46]. However, the relevance of these isoforms to disease is, as of yet, poorly understood.

1.5 Mechanisms of TDP43 Autoregulation

Given that TDP43 recognizes UG-rich sequences present within approximately one third of all transcribed genes[44,47,48], it is uniquely able to influence the processing of hundreds to thousands of transcripts. In keeping with these fundamental functions, the level and localization of TDP43 are tightly regulated and critical for cell health. TDP43 knockout mice die early in embryogenesis, and partial or conditional knockout animals exhibit neurodegeneration and behavioral deficits that correlate with the neuroanatomical pattern of TDP43 ablation[49–52]. Additionally, sustained TDP43 overexpression results in neurodegeneration in primary neuron[53], mouse[54,55], rat[56,57], Drosophila[58,59], zebrafish[60,61], and primate models[62,63], providing convincing evidence that too little or too much TDP43 is lethal.

Despite the observed sensitivity of neurons and other cell types to long-term changes in TDP43 protein levels, TDP43 expression and localization are dynamically regulated in the short-term by physical injury and other cellular stressors[64–66]. This pattern of expression suggests that TDP43 may be important for orchestrating the response to acute injury and eventual recovery. However, even relatively minor (~2-fold) persistent changes in TDP43 levels are sufficient to drive neurodegeneration[53,67–69], indicative of a coping response that over time becomes ineffective and eventually detrimental to cell health.

Similar to systems employed by related RNA-binding proteins, TDP43 regulates its own expression through an intricate negative feedback loop. At high levels, TDP43 recognizes sequences within the 3' untranslated region (UTR) of its own transcript (the TDP43 binding region, or TDPBR)[43,70], triggering alternative splicing within the 3' UTR[44,47], mRNA destabilization, and reduced protein expression[43,44,71]. Two separate mechanisms may account for this destabilization (Figure 1.2).

In the first, association of TDP43 with the TDPBR induces the removal of two alternative introns (6 and 7) within the last exon of the *TARDBP* mRNA transcript[44,72]. These splicing events create perceived exon-exon junctions (EEJs) with subsequent deposition of exon-junction complexes (EJCs), structures composed of eukaryotic initiation factor 4A-III, Magoh, Y14, UPF2 and UPF3. During the process of translation, scanning ribosomes typically displace EJCs at EEJs upstream of a stop codon. Translation is stalled when the ribosome encounters a stop codon, allowing association of the SURF complex (SMG1, UPF1, and eRF1 and 2) with the ribosome. When an EJC is present >50 nt downstream of the stop codon, factors within the EJC (i.e. UPF2) may interact with UPF1 in the SURF complex, triggering UPF1 phosphorylation and nonsense-mediated mRNA decay (NMD)[73,74]. In support of this model, knockdown of UPF1 — an essential NMD factor[74–76] — increased the expression of constructs carrying the *TARDBP* 3' UTR, while exogenous TDP43 reduced their expression[44,77].

This mechanism of autoregulation by RNA-binding proteins is not unique to TDP43, and forms the basis for a cascade labeled **r**egulated **u**nproductive **s**plicing and **t**ranslation (RUST) that is also utilized by the splicing factors PTB and SC35[78–82]. Like TDP43, these proteins recognize sequences present within the 3' UTR of their respective transcripts, resulting in splicing and EJC deposition downstream of the canonical stop codon. This, in turn, causes RNA destabilization via

NMD, and an overall reduction in protein levels. An analogous mechanism is responsible for the regulation of FUS, a nuclear RNA-binding protein whose cytoplasmic mislocalization and accumulation are implicated in ALS, much like TDP43[83–86]. FUS and TDP43 share basic structural and functional elements, including a glycine-rich low complexity domain that harbors ALS-associated mutations. FUS also binds its own transcript, resulting in exclusion of exon 7 and a shift in the reading frame[87]. This shift uncovers a premature stop codon in exon 8, leading to destabilization of the alternatively-spliced *FUS* mRNA via NMD. Furthermore, disease-associated mutations in *FUS*[87] and *TARDBP*[72] may impair effective autoregulation of these RNA-binding proteins, resulting in their accumulation and downstream toxicity (see below).

Nevertheless, alternatively-spliced *TARDBP* mRNA isoforms and predicted NMD substrates have been difficult to identify and measure, and additional studies suggest that TDP43 autoregulation operates by a separate mechanism. In the second model, TDP43-mediated splicing within the *TARDBP* 3' UTR removes the primary mRNA polyadenylation site (pA1) present within intron 7[43,88]. Transcripts that utilize the remaining polyadenylation sites pA2 and pA4 are preferentially retained in the nucleus[88] and degraded by the RNA exosome[43]. Genetic ablation of exosome components Rrp6 and Rrp44 is sufficient to increase exogenous *TARDBP* mRNA levels and protein production, implying that the RNA exosome is indeed responsible for degrading the overexpressed *TARDBP* minigene. However, more recent evidence suggests that differential polyadenylation cannot fully explain TDP43 autoregulation in this model[89]. Rather, TDP43-induced splicing of intron 7 within the *TARDBP* 3' UTR destabilizes the transcript, reduces nuclear export, and decreases protein production. Artificial mutations that enhance intron 7 splicing promote *TARDBP* destabilization, while cDNA and other transcripts that intrinsically lack intron 7 escape autoregulation and are constitutively expressed at high levels[89]. This suggests that

spliceosome assembly and intron 7 splicing is a key event in TDP43 autoregulation, but whether this process participates in the regulation of endogenous TDP43 levels is unclear, and further studies are required to fully elucidate its contribution.

1.6 Disruption of TDP43 Autoregulation in ALS

Regardless of whether TDP43 mRNA is destabilized by NMD or degraded by the exosome following nuclear retention, interruption of this autoregulatory process likely has severe consequences for cell health. Five disease-associated mutations have been identified within the *TARDBP* 3' UTR[90], which may block binding of TDP43 to its own transcript and subsequent alternative splicing. At least one of these mutations is associated with a steady-state increase in *TARDBP* mRNA levels, supporting the notion of disrupted autoregulation as an underlying factor leading to TDP43 accumulation and disease[91]. The majority of ALS-associated *TARDBP* mutations lie within the carboxy-terminal glycine rich domain of the protein[92], and although the precise mechanism remains unclear, several studies have suggested that these pathogenic mutations enhance cytoplasmic TDP43 mislocalization and aggregation[53,69,93,94] and stabilize cytoplasmic TDP43[95,96]. By increasing the proportion of cytoplasmic TDP43, these changes would be expected to reduce autoregulation, resulting in elevated TDP43 production. Eventually, this vicious cycle may culminate in cytoplasmic TDP43 deposition, nuclear TDP43 clearance and neurodegeneration[97].

Mutations in genes other than *TARDBP* may inhibit TDP43 autoregulation by affecting nucleocytoplasmic transport. The most prevalent mutation responsible for ALS, hexanucleotide (G_4C_2) expansions in the *C9orf72* gene, may block nuclear protein import through one or more related mechanisms: repeat (G_4C_2)-containing RNA may sequester essential transport factors (i.e.

RanGAP)[98], or dipeptide repeat proteins produced by repeat-associated non-AUG (RAN) translation of the repeat RNA might directly clog the nuclear pore[99]. Additional evidence suggests that cytoplasmic protein aggregates may universally impair nucleocytoplasmic transport in neurodegenerative conditions[100], further facilitating the cytoplasmic deposition of proteins such as TDP43 and FUS that typically participate in nucleocytoplasmic shuttling. Because splicing takes place within the nucleus, cytoplasmic retention of FUS and TDP43 would interfere with the normal autoregulation process, ultimately increasing mRNA stability and protein production. Therefore, cytoplasmic sequestration of the proteins or inefficient nuclear import would be sufficient to inhibit autoregulation, accelerating the formation of TDP43 or FUS cytoplasmic inclusions that are characteristic of ALS[97].

1.7 Downstream Consequences of Failed TDP43 Autoregulation: TDP43 Nuclear Exclusion and Cytoplasmic Accumulation

Disruption of TDP43 autoregulation influences both the protein level and localization of TDP43, resulting in cytoplasmic deposition and nuclear clearance characteristic of the majority of ALS patients. However, how exactly TDP43 dysfunction contributes to disease pathogenesis remains unclear. Here, we will explore 1.) the consequences of disrupted normal protein function and 2.) the potential gain of toxic function that occurs in conjunction with TDP43 pathology.

1.7.1 Disrupted TDP43 Autoregulation Interrupts Normal Protein Function

Given TDP43's crucial functions in RNA processing, its dysregulation leads to abnormalities in alternative mRNA splicing, non-coding RNAs, miRNA biogenesis, and the dynamics of RNA-rich granules (Figure 1.3).

1.7.1.1 Alternative Splicing

Alternative splicing is the differential inclusion or exclusion of exons within mature transcripts, enabling the expression of multiple RNA and protein isoforms from a single gene. Between 92 and 94% of all mRNAs in the human genome are alternatively spliced[101], and the brain expresses more alternatively spliced mRNAs than any other organ[102,103]. Because changes in the splicing environment determine which isoforms are produced[104,105], alternative splicing can regulate gene expression by creating transcripts that are more or less stable. In fact, an estimated 33% of alternatively-spliced transcripts contain premature termination codons that mark them as substrates for NMD[106]. Thus, NMD is not simply a mechanism for degrading abnormal or mutated transcripts, but also represents an active pathway regulating the stability of alternatively-spliced transcripts. Alternative splicing therefore represents an effective and rapid means of regulating gene expression via changes in RNA stability, without the need to revert to transcription. As mentioned above, RUST is an NMD-related mechanism utilized by RNA-binding proteins to dynamically and quickly modulate their own expression; signal transduction by inflammatory cytokines likewise affects gene expression via changes in splicing and RNA stability[107,108].

In addition to regulating the splicing of its own mRNA, TDP43 is crucial for the alternative splicing of hundreds of other transcripts[44,47,84,109]. It interacts strongly with several splicing factors[110], and loss of TDP43 causes widespread changes in alternative splicing[44,47] including many transcripts that are critical for neuronal viability[44,84,111]. ALS-associated *TARDBP* mutations can likewise alter alternative splicing and subsequent gene expression[109,112].

Alternatively spliced transcripts can also be targeted for decay if they include mutations that create novel splice sites. This can lead to the inclusion of unannotated or “cryptic” exons and

the production of faulty transcripts. Recent work suggests that TDP43 actively suppresses unannotated exon splicing events[113,114]. Its depletion results in a widespread increase in cryptic exon splicing, and the inclusion of these exons typically leads to NMD[115]. Many of these events are specific to neurons[116], suggesting that the disruption of TDP43-mediated cryptic exon regulation may directly contribute to neurodegeneration. Furthermore, the unannotated exons affected by TDP43 are distinct in murine and human cells[113], indicating species-specific differences in TDP43 function that may predispose to mechanisms of neurodegeneration unique to humans. TDP43 is not the only RNA-binding protein that modulates exon inclusion — RBM17, PTBP1 and PTBP2 also repress the inclusion of unannotated exons[113,114]. Like TDP43, each of these factors is essential for neuronal development and their loss results in neurodegeneration[114,117–119], implying that neurons are particularly susceptible to the abnormal inclusion of unannotated exons.

1.7.1.2 Non-Coding RNAs

Though attention is often focused on the 1-2% of transcripts that encode protein, the vast majority of the genome is transcribed as non-protein-coding RNAs (ncRNAs)[120]. These transcripts are loosely categorized as short or long non-coding RNAs (lncRNAs), and the latter act by regulating gene expression in a variety of ways. These include, but are not limited to: the sequestration[121], competition[122], or altered localization of transcription factors[123]; transcriptional coactivation[124] or corepression[125]; alternative mRNA splicing[126]; mRNA transport and stability[127]; and modulation of translation[128]. lncRNAs serve crucial functions in development and disease, and also help scaffold membraneless organelles such as nuclear speckles and paraspeckles[129] that are important sites of RNA processing and modification[130].

Both TDP43 and FUS recognize lncRNAs[47,84,131], including *gadd7*[132], MALAT1[133], and NEAT1_2[134], via UG-rich binding sites[47,84]. The abundance of many lncRNAs is altered in response to TDP43 knockdown in murine models of ALS[44] and in human post-mortem tissue[47]. Thus, TDP43 deposition in ALS likely has profound consequences for lncRNA expression and function. However, further studies are required to determine how TDP43 pathology influences lncRNA-related processes, and whether TDP43-mediated impairment of lncRNA contributes significantly to neurodegeneration in ALS.

1.7.1.3 miRNA Biogenesis

MicroRNAs (miRNAs) are small, non-coding RNAs that base-pair with complementary sequences within mRNA transcripts to trigger their decay and/or translational repression. These 20-25 nt RNAs are produced from an RNA precursor (pri-miRNA) that forms a hairpin loop shortly after transcription[135,136]. The enzyme Drosha then cleaves the hairpin from the rest of the transcript[137,138], and the resulting molecule (pre-miRNA) is exported to the cytoplasm[139]. There, the enzyme Dicer cuts away the looped end[140], leaving a duplex of two short, complementary RNA strands. The two strands dissociate and the mature miRNA associates with the RNA-induced silencing complex (RISC), which assists in orienting the miRNA with its mRNA target, repressing translation of the target transcript and triggering its degradation.

TDP43 promotes miRNA biogenesis through a direct association with pri-miRNA, pre-miRNA, and both Drosha and Dicer[141]. In so doing, TDP43 regulates the formation of key miRNAs that are essential for neuronal development, activity and survival[141–145]. FUS also interacts with Drosha and pri-miRNA in neurons, suggesting that it plays a similar role in neuronal miRNA biogenesis[146].

Concordant with TDP43 pathology, the expression of several TDP43-associated miRNAs were altered in the CSF of sporadic ALS patients, compared to healthy controls[147,148]. Similar changes in miRNA levels were detected in transgenic mutant SOD1 mouse spinal cord and human ALS monocytes, but not fibroblasts from ALS patients[149,150]. Human neurons carrying *TARDBP* mutations exhibited reduced levels of miR-9 and the immature pri-miRNA precursor pri-miR-9-2[145]. Knockdown of endogenous *TARDBP* in control neurons reproduced these deficits, suggesting that TDP43 actively participates in miR-9 biogenesis, and that disease-associated *TARDBP* mutations inhibit this function. One of the predicted targets of miRNAs disrupted in ALS tissues is EIF2/AGO4[148], a component of RISC that participates in miRNA-mediated RNA degradation[151]. Thus, abnormal miRNA biogenesis triggered by TDP43 dysfunction in ALS may have direct and indirect consequences for the maintenance of RNA stability. Further, since each individual miRNA can regulate the stability and translation of many downstream mRNA targets, the potential implications of even minor abnormalities in miRNA biogenesis are considerable[152].

1.7.1.4 Stress Granule Dynamics

Cells undergo a wide range of molecular changes in response to environmental stressors, including the inhibition of conventional translation[153,154] and the formation of stress granules (SGs), cytoplasmic ribonucleoprotein particles rich in mRNA, RNA-binding proteins, and stalled translation initiation complexes[155–157]. TDP43 is one of several RNA-binding proteins that localize to SGs in response to various conditions[158–161]. Although it is not essential for SG formation *per se*, changes in TDP43 levels or localization affect SG dynamics. For instance, *TARDBP* knockdown slows SG formation[158,161], while expression of ALS-associated mutant TDP43 accelerates SG formation and results in larger SGs than wild-type TDP43 overexpression[159,160].

Based on its ability to recognize thousands of GU-rich transcripts, it is possible that excess TDP43 inclusion within SGs enables broad mRNA sequestration, shifting transcripts from actively translating polysomes to the relatively inert SGs. Conversely, SG-localized TDP43 may also bind to and prevent the degradation of RNAs that would have otherwise been degraded through association with components of processing (P)-bodies, including decapping proteins and exonucleases. Therefore, cytoplasmic TDP43 accumulation within normal or abnormal SGs in ALS might effectively increase mRNA stability without causing a reciprocal increase in mRNA translation. Nevertheless, any potential RNA stabilizing effect of TDP43 deposition is likely to be outweighed by the substantive TDP43-dependent changes in alternative splicing, unannotated exon inclusion, and miRNA biogenesis that collectively act to destabilize RNA.

1.7.1.5 RNA Transport Granules

Localized translation of mRNA is a common mechanism for regulating protein expression in specific regions of the cell. This is of particular importance in highly compartmentalized cells such as neurons, in which local translation is essential for synaptic plasticity[162], neurotransmitter production[163], axon guidance, and recovery from injury[164]. mRNAs are transported in granules comprised primarily of RNA-binding proteins[165] that stabilize and translationally repress[166,167] their cargo. TDP43 colocalizes with mRNA and related RNA-binding proteins in transport granules that undergo bidirectional, microtubule-dependent transport[168,169], suggesting that TDP43 acts as a neuronal mRNA transport factor[170]. Disease-associated *TARDBP* mutations impair the motility of TDP43-positive axonal granules[168], and the overexpression of TDP43 C-terminal fragments sequester components of transport granules such as HuD[169]. Taken together with evidence showing that wild-type or mutant TDP43 overexpression impairs axon outgrowth[169],

these observations imply that TDP43-dependent dysregulation of mRNA transport and local protein synthesis may contribute to axon degeneration in ALS.

1.7.2 TDP43 Inclusions and Gain of Function Toxicity

The role of TDP43 aggregates in neurodegeneration is not well understood. In contrast to the hypothesis that neurodegeneration in ALS results from the loss of normal TDP43 function, others have proposed that cytosolic aggregates gain and exert novel toxic functions. As previously described, TDP43 in cytosolic aggregates undergoes several histopathologic changes rendering it insoluble, ubiquitinated[17], and hyperphosphorylated[22,23]. Moreover, a prominent species within these inclusions are 25-35 kDa TDP43 C-terminal fragments[171]. The C-terminal domain harbors the majority of ALS-associated mutations and phosphorylation sites, and this domain alone is highly aggregation prone and sufficient to induce cell death in some contexts[93,172]. Moreover, there is some evidence to suggest that this C-terminal domain is capable of forming amyloid-like fibrils that sequester critical cellular components and disrupt their normal function[93,173,174].

Altered aggregation propensity is not limited to the C-terminus; the RRMs play a role in protein aggregation[175] that is perhaps exacerbated by RNA binding[176]. The N-terminus also promotes TDP43 oligomerization[177] and may contribute to the formation of TDP43 aggregates[27,28,178]. However, there is limited evidence directly linking inclusions to neurodegeneration[179]. Experimental models have demonstrated that TDP43 aggregates are not necessary for TDP43-mediated toxicity[180], and although that does not preclude their involvement in neuron dysfunction and death, existing evidence suggests that TDP43 mislocalization and expression levels are more significant contributors to neurodegeneration in ALS[53,69,109].

1.8 Summary and Aims of the Dissertation

ALS is a devastating neurodegenerative disorder, and its heterogeneous biochemical, genetic, and clinical features complicate the identification of therapeutic targets. However, the cytoplasmic mislocalization and accumulation of TDP43, a nuclear RNA-binding protein integrally involved in RNA metabolism, is observed in the vast majority of individuals with ALS. TDP43 is an essential protein involved in several RNA processing events, including splicing, translation, and degradation. As such, TDP43 levels and localization are tightly regulated, where even small perturbations in abundance and localization are sufficient to induce cell death. However, the initiating events that drive TDP43 pathology remain unclear, as do the contributions of TDP43 dysregulation to ALS pathogenesis.

The primary goal of my dissertation is to identify and understand convergent mechanisms that lead to neurodegeneration in ALS. My central hypothesis is that the dysregulation of TDP43 leads to neurodegeneration via a combination of TDP43 loss-of-function and gain-of-function toxicity. To address this, my dissertation is divided into the following aims: 1.) explore other features common to the majority of ALS cases and their relation to TDP43 dysregulation, and 2.) identify how changes in TDP43 metabolism contribute to TDP43 pathology. Chapter 2 provides a broad examination of RNA metabolism both in the context of healthy cells and neurodegenerative disease. Chapter 3 examines how another hallmark feature of ALS, hyperexcitability, drives the formation of truncated TDP43 isoforms that recapitulate and perhaps exacerbate TDP43 pathology, and Appendix A describes progress made in using an induced pluripotent stem cell-derived neuron model system to track TDP43 synthesis and degradation at the endogenous level using longitudinal fluorescence microscopy, which is detailed further in Appendix B. Appendix C

describes the establishment of a medium-throughput screening to identify novel therapeutics for ALS. Finally, Chapter 4 concludes the dissertation and describes future directions.

1.9 Acknowledgements

Funding for this work was provided by the National Institutes of Health / National Institute for Neurological Disorders and Stroke (NIH/NINDS) R01NS097542 (SB), and Ann Arbor Active Against ALS (KW).

References

1. Bruijn, L. I., Miller, T. M. & Cleveland, D. W. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu. Rev. Neurosci.* **27**, 723–749 (2004).
2. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* **17**, 17–23 (2014).
3. Taylor, J. P., Brown, R. H., Jr & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).
4. Petrov, D., Mansfield, C., Moussy, A. & Hermine, O. ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment? *Front. Aging Neurosci.* **9**, 68 (2017).
5. Miller, R. G., Mitchell, J. D. & Moore, D. H. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst. Rev.* CD001447 (2012). doi:10.1002/14651858.CD001447.pub3
6. Yoshino, H. & Kimura, A. Investigation of the therapeutic effects of edaravone, a free radical scavenger, on amyotrophic lateral sclerosis (Phase II study). *Amyotroph. Lateral Scler.* **7**, 241–245 (2006).
7. Abe, K. *et al.* Confirmatory double-blind, parallel-group, placebo-controlled study of efficacy and safety of edaravone (MCI-186) in amyotrophic lateral sclerosis patients. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **15**, 610–617 (2014).
8. Kapeli, K., Martinez, F. J. & Yeo, G. W. Genetic mutations in RNA-binding proteins and their roles in ALS. *Hum. Genet.* (2017). doi:10.1007/s00439-017-1830-7
9. Neumann, M. Molecular neuropathology of TDP-43 proteinopathies. *Int. J. Mol. Sci.* **10**, 232–246 (2009).
10. Murray, M. E. *et al.* Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. *Acta Neuropathol.* **122**, 673–690 (2011).
11. Seilhean, D. *et al.* Accumulation of TDP-43 and α-actin in an amyotrophic lateral sclerosis patient with the K17I ANG mutation. *Acta Neuropathol.* **118**, 561–573 (2009).
12. Van Mossevelde, S. *et al.* Clinical features of TBK1 carriers compared with C9orf72, GRN and non-mutation carriers in a Belgian cohort. *Brain* **139**, 452–467 (2016).
13. Smith, B. N. *et al.* Novel mutations support a role for Profilin 1 in the pathogenesis of ALS. *Neurobiol. Aging* **36**, 1602.e17–27 (2015).
14. Deng, H.-X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).
15. Johnson, J. O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).
16. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).
17. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
18. Mackenzie, I. R. A. *et al.* Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Ann. Neurol.* **61**, 427–434 (2007).
19. Leigh, P. N. *et al.* Ubiquitin-immunoreactive intraneuronal inclusions in amyotrophic lateral sclerosis. Morphology, distribution, and specificity. *Brain* **114 (Pt 2)**, 775–788 (1991).

20. Lowe, J. New pathological findings in amyotrophic lateral sclerosis. *J. Neurol. Sci.* **124 Suppl**, 38–51 (1994).
21. Brettschneider, J. *et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Annals of Neurology* **74**, 20–38 (2013).
22. Hasegawa, M. *et al.* Phosphorylated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **64**, 60–70 (2008).
23. Inukai, Y. *et al.* Abnormal phosphorylation of Ser409/410 of TDP-43 in FTLD-U and ALS. *FEBS Lett.* **582**, 2899–2904 (2008).
24. Igaz, L. M. *et al.* Enrichment of C-Terminal Fragments in TAR DNA-Binding Protein-43 Cytoplasmic Inclusions in Brain but not in Spinal Cord of Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *The American Journal of Pathology* **173**, 182–194 (2008).
25. Ou, S. H., Wu, F., Harrich, D., García-Martínez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J. Virol.* **69**, 3584–3596 (1995).
26. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
27. Shiina, Y., Arima, K., Tabunoki, H. & Satoh, J.-I. TDP-43 dimerizes in human cells in culture. *Cell. Mol. Neurobiol.* **30**, 641–652 (2010).
28. Zhang, Y.-J. *et al.* The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112–3122 (2013).
29. Winton, M. J. *et al.* Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *J. Biol. Chem.* **283**, 13302–13309 (2008).
30. Ayala, Y. M. *et al.* Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778–3785 (2008).
31. Kuo, P.-H., Doudeva, L. G., Wang, Y.-T., Shen, C.-K. J. & Yuan, H. S. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* **37**, 1799–1808 (2009).
32. Lukavsky, P. J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449 (2013).
33. Kuo, P.-H. *et al.* The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Research* **42**, 4712–4722 (2014).
34. Ayala, Y. M. *et al.* Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol.* **348**, 575–588 (2005).
35. Buratti, E. & Baralle, F. E. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator ofCFTR Exon 9. *J. Biol. Chem.* **276**, 36337–36343 (2001).
36. D'Ambrogio, A. *et al.* Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Res.* **37**, 4116–4126 (2009).
37. Buratti, E. *et al.* TDP-43 binds heterogeneous nuclear ribonucleoprotein a/b through its c-terminal tail an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J. Biol. Chem.* **280**, 37572–37584 (2005).

38. Deshaies, J.-E. *et al.* TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis. *Brain* **141**, 1320–1333 (2018).
39. Santamaria, N., Alhothali, M., Alfonso, M. H., Breydo, L. & Uversky, V. N. Intrinsic disorder in proteins involved in amyotrophic lateral sclerosis. *Cell. Mol. Life Sci.* **74**, 1297–1318 (2017).
40. Franzmann, T. & Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* (2018). doi:10.1074/jbc.TM118.001190
41. Wang, H.-Y., Wang, I.-F., Bose, J. & Shen, C.-K. J. Structural diversity and functional implications of the eukaryotic TDP gene family. *Genomics* **83**, 130–139 (2004).
42. D'Alton, S., Altshuler, M. & Lewis, J. Studies of alternative isoforms provide insight into TDP-43 autoregulation and pathogenesis. *RNA* **21**, 1419–1432 (2015).
43. Ayala, Y. M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).
44. 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).
45. Seyfried, N. T. *et al.* Multiplex SILAC analysis of a cellular TDP-43 proteinopathy model reveals protein inclusions associated with SUMOylation and diverse polyubiquitin chains. *Mol. Cell. Proteomics* **9**, 705–718 (2010).
46. Dammer, E. B. *et al.* Coaggregation of RNA-binding proteins in a model of TDP-43 proteinopathy with selective RGG motif methylation and a role for RRM1 ubiquitination. *PLoS One* **7**, e38658 (2012).
47. Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
48. Sephton, C. F. *et al.* Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J. Biol. Chem.* **286**, 1204–1215 (2011).
49. Kraemer, B. C. *et al.* Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis. *Acta Neuropathol.* **119**, 409–419 (2010).
50. Wu, L.-S., Cheng, W.-C. & Shen, C.-K. J. Targeted depletion of TDP-43 expression in the spinal cord motor neurons leads to the development of amyotrophic lateral sclerosis-like phenotypes in mice. *J. Biol. Chem.* **287**, 27335–27344 (2012).
51. Iguchi, Y. *et al.* Loss of TDP-43 causes age-dependent progressive motor neuron degeneration. *Brain* **136**, 1371–1382 (2013).
52. Sephton, C. F. *et al.* TDP-43 is a developmentally regulated protein essential for early embryonic development. *J. Biol. Chem.* **285**, 6826–6834 (2010).
53. Barmada, S. J. *et al.* Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
54. Swarup, V. *et al.* Pathological hallmarks of amyotrophic lateral sclerosis/frontotemporal lobar degeneration in transgenic mice produced with TDP-43 genomic fragments. *Brain* **134**, 2610–2626 (2011).
55. Wils, H. *et al.* TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3858–3863 (2010).
56. Dayton, R. D. *et al.* Selective forelimb impairment in rats expressing a pathological TDP-43 25 kDa C-terminal fragment to mimic amyotrophic lateral sclerosis. *Mol. Ther.* **21**, 1324–

1334 (2013).
57. Tatom, J. B. *et al.* Mimicking aspects of frontotemporal lobar degeneration and Lou Gehrig's disease in rats via TDP-43 overexpression. *Mol. Ther.* **17**, 607–613 (2009).
58. Voigt, A. *et al.* TDP-43-mediated neuron loss in vivo requires RNA-binding activity. *PLoS One* **5**, e12247 (2010).
59. Li, Y. *et al.* A Drosophila model for TDP-43 proteinopathy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3169–3174 (2010).
60. Kabashi, E. *et al.* Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum. Mol. Genet.* **19**, 671–683 (2010).
61. Schmid, B. *et al.* Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4986–4991 (2013).
62. Uchida, A. *et al.* Non-human primate model of amyotrophic lateral sclerosis with cytoplasmic mislocalization of TDP-43. *Brain* **135**, 833–846 (2012).
63. Jackson, K. L. *et al.* Initial gene vector dosing for studying symptomatology of amyotrophic lateral sclerosis in non-human primates. *J. Med. Primatol.* **44**, 66–75 (2015).
64. Moisse, K. *et al.* Divergent patterns of cytosolic TDP-43 and neuronal progranulin expression following axotomy: implications for TDP-43 in the physiological response to neuronal injury. *Brain Res.* **1249**, 202–211 (2009).
65. Swarup, V., Audet, J.-N., Phaneuf, D., Kriz, J. & Julien, J.-P. Abnormal regenerative responses and impaired axonal outgrowth after nerve crush in TDP-43 transgenic mouse models of amyotrophic lateral sclerosis. *J. Neurosci.* **32**, 18186–18195 (2012).
66. Johnson, V. E., Stewart, W., Trojanowski, J. Q. & Smith, D. H. Acute and chronically increased immunoreactivity to phosphorylation-independent but not pathological TDP-43 after a single traumatic brain injury in humans. *Acta Neuropathol.* **122**, 715–726 (2011).
67. Janssens, J. *et al.* Overexpression of ALS-associated p.M337V human TDP-43 in mice worsens disease features compared to wild-type human TDP-43 mice. *Mol. Neurobiol.* **48**, 22–35 (2013).
68. Wegorzewska, I. & Baloh, R. H. TDP-43-based animal models of neurodegeneration: new insights into ALS pathology and pathophysiology. *Neurodegener. Dis.* **8**, 262–274 (2011).
69. Barmada, S. J. *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat. Chem. Biol.* **10**, 677–685 (2014).
70. Bhardwaj, A., Myers, M. P., Buratti, E. & Baralle, F. E. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* **41**, 5062–5074 (2013).
71. Xu, Y.-F. *et al.* Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J. Neurosci.* **30**, 10851–10859 (2010).
72. Koyama, A. *et al.* Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic Acids Res.* **44**, 5820–5836 (2016).
73. Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W. & Kulozik, A. E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747 (2008).
74. Popp, M. W.-L. & Maquat, L. E. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu. Rev. Genet.* **47**, 139–165 (2013).
75. Sun, X., Perlick, H. A., Dietz, H. C. & Maquat, L. E. A mutated human homologue to yeast

Upf1 protein has a dominant-negative effect on the decay of nonsensecontaining mRNAs in mammalian cells. *Proceedings of the National Academy of Sciences* **95**, 10009–10014 (1998).
76. Medghalchi, S. M. *et al.* Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* **10**, 99–105 (2001).
77. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
78. Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. J. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* **13**, 91–100 (2004).
79. Sureau, A., Gattoni, R., Dooghe, Y., Stévenin, J. & Soret, J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J.* **20**, 1785–1796 (2001).
80. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **21**, 708–718 (2007).
81. Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C. & Brenner, S. E. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**, 926–929 (2007).
82. Dredge, B. K., Stefani, G., Engelhard, C. C. & Darnell, R. B. Nova autoregulation reveals dual functions in neuronal splicing. *EMBO J.* **24**, 1608–1620 (2005).
83. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: the FUS about TDP-43. *Cell* **136**, 1001–1004 (2009).
84. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488–1497 (2012).
85. Kwiatkowski, T. J., Jr *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).
86. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).
87. Zhou, Y., Liu, S., Liu, G., Oztürk, A. & Hicks, G. G. ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. *PLoS Genet.* **9**, e1003895 (2013).
88. Avendaño-Vázquez, S. E. *et al.* Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev.* **26**, 1679–1684 (2012).
89. Bembich, S. *et al.* Predominance of spliceosomal complex formation over polyadenylation site selection in TDP-43 autoregulation. *Nucleic Acids Res.* **42**, 3362–3371 (2014).
90. Pesiridis, G. S., Lee, V. M.-Y. & Trojanowski, J. Q. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **18**, R156–62 (2009).
91. Gitcho, M. A. *et al.* TARDBP 3′-UTR variant in autopsy-confirmed frontotemporal lobar degeneration with TDP-43 proteinopathy. *Acta Neuropathol.* **118**, 633 (2009).
92. Barmada, S. J. & Finkbeiner, S. Pathogenic TARDBP mutations in amyotrophic lateral sclerosis and frontotemporal dementia: disease-associated pathways. *Rev. Neurosci.* **21**, 251–272 (2010).
93. Guo, W. *et al.* An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat. Struct. Mol. Biol.* **18**, 822–830 (2011).

94. Mutihac, R. *et al.* TARDBP pathogenic mutations increase cytoplasmic translocation of TDP-43 and cause reduction of endoplasmic reticulum Ca^{2+} signaling in motor neurons. *Neurobiol. Dis.* **75**, 64–77 (2015).
95. Austin, J. A. *et al.* Disease causing mutants of TDP-43 nucleic acid binding domains are resistant to aggregation and have increased stability and half-life. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4309–4314 (2014).
96. Ling, S.-C. *et al.* ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13318–13323 (2010).
97. Lee, E. B., Lee, V. M.-Y. & Trojanowski, J. Q. Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nat. Rev. Neurosci.* **13**, 38–50 (2011).
98. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56–61 (2015).
99. Shi, K. Y. *et al.* Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1111–E1117 (2017).
100. Woerner, A. C. *et al.* Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science* **351**, 173–176 (2016).
101. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
102. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
103. Yeo, G., Holste, D., Kreiman, G. & Burge, C. B. Variation in alternative splicing across human tissues. *Genome Biol.* **5**, R74 (2004).
104. Weg-Remers, S., Ponta, H., Herrlich, P. & König, H. Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.* **20**, 4194–4203 (2001).
105. van der Houven van Oordt, W. *et al.* The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J. Cell Biol.* **149**, 307–316 (2000).
106. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 189–192 (2003).
107. Shakola, F., Suri, P. & Ruggiu, M. Splicing Regulation of Pro-Inflammatory Cytokines and Chemokines: At the Interface of the Neuroendocrine and Immune Systems. *Biomolecules* **5**, 2073–2100 (2015).
108. Paulsen, M. T. *et al.* Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2240–2245 (2013).
109. Arnold, E. S. *et al.* 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. U. S. A.* **110**, E736–45 (2013).
110. Freibaum, B. D., Chitta, R. K., High, A. A. & Taylor, J. P. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J. Proteome Res.* **9**, 1104–1120 (2010).
111. Chang, J.-C., Hazelett, D. J., Stewart, J. A. & Morton, D. B. Motor neuron expression of the voltage-gated calcium channel cacophony restores locomotion defects in a Drosophila,

TDP-43 loss of function model of ALS. *Brain Res.* **1584**, 39–51 (2014).
112. Highley, J. R. *et al.* Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones. *Neuropathol. Appl. Neurobiol.* **40**, 670–685 (2014).
113. 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).
114. Tan, Q. *et al.* Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum. Mol. Genet.* **25**, 5083–5093 (2016).
115. Humphrey, J., Emmett, W., Fratta, P., Isaacs, A. M. & Plagnol, V. Quantitative analysis of cryptic splicing associated with TDP-43 depletion. *BMC Med. Genomics* **10**, 38 (2017).
116. Jeong, Y. H. *et al.* Tdp-43 cryptic exons are highly variable between cell types. *Mol. Neurodegener.* **12**, 13 (2017).
117. Li, Q. *et al.* The splicing regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. *Elife* **3**, e01201 (2014).
118. Suckale, J. *et al.* PTBP1 is required for embryonic development before gastrulation. *PLoS One* **6**, e16992 (2011).
119. Ling, J. P. *et al.* PTBP1 and PTBP2 Repress Nonconserved Cryptic Exons. *Cell Rep.* **17**, 104–113 (2016).
120. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **12**, 861–874 (2011).
121. Hung, T. *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* **43**, 621–629 (2011).
122. Kino, T., Hurt, D. E., Ichijo, T., Nader, N. & Chrousos, G. P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* **3**, ra8 (2010).
123. Liu, Z. *et al.* The kinase LRRK2 is a regulator of the transcription factor NFAT that modulates the severity of inflammatory bowel disease. *Nat. Immunol.* **12**, 1063–1070 (2011).
124. Hubé, F., Velasco, G., Rollin, J., Furling, D. & Francastel, C. Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. *Nucleic Acids Res.* **39**, 513–525 (2011).
125. Wang, X. *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **454**, 126–130 (2008).
126. Annilo, T., Kepp, K. & Laan, M. Natural antisense transcript of natriuretic peptide precursor A (NPPA): structural organization and modulation of NPPA expression. *BMC Mol. Biol.* **10**, 81 (2009).
127. Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284–288 (2011).
128. Ebralidze, A. K. *et al.* PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev.* **22**, 2085–2092 (2008).
129. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long noncoding RNAs: past, present, and future. *Genetics* **193**, 651–669 (2013).
130. Galganski, L., Urbanek, M. O. & Krzyzosiak, W. J. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res.* **45**, 10350–10368 (2017).
131. Hoell, J. I. *et al.* RNA targets of wild-type and mutant FET family proteins. *Nat. Struct.*

Mol. Biol. **18**, 1428–1431 (2011).
132. Liu, X., Li, D., Zhang, W., Guo, M. & Zhan, Q. Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. *EMBO J.* **31**, 4415–4427 (2012).
133. Guo, F. *et al.* Regulation of MALAT1 expression by TDP43 controls the migration and invasion of non-small cell lung cancer cells in vitro. *Biochem. Biophys. Res. Commun.* **465**, 293–298 (2015).
134. Nishimoto, Y. *et al.* The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol. Brain* **6**, 31 (2013).
135. Cai, X., Hagedorn, C. H. & Cullen, B. R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957–1966 (2004).
136. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
137. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
138. Gregory, R. I., Chendrimada, T. P. & Shiekhattar, R. MicroRNA biogenesis: isolation and characterization of the microprocessor complex. *Methods Mol. Biol.* **342**, 33–47 (2006).
139. Murchison, E. P. & Hannon, G. J. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr. Opin. Cell Biol.* **16**, 223–229 (2004).
140. Lund, E. & Dahlberg, J. E. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 59–66 (2006).
141. Kawahara, Y. & Mieda-Sato, A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3347–3352 (2012).
142. Fan, Z., Chen, X. & Chen, R. Transcriptome-wide analysis of TDP-43 binding small RNAs identifies miR-NID1 (miR-8485), a novel miRNA that represses NRXN1 expression. *Genomics* **103**, 76–82 (2014).
143. Buratti, E. *et al.* Nuclear factor TDP-43 can affect selected microRNA levels. *FEBS J.* **277**, 2268–2281 (2010).
144. Gascon, E. & Gao, F.-B. The Emerging Roles of MicroRNAs in the Pathogenesis of Frontotemporal Dementia–Amyotrophic Lateral Sclerosis (FTD-ALS) Spectrum Disorders. *J. Neurogenet.* **28**, 30–40 (2014).
145. Zhang, Z. *et al.* Downregulation of microRNA-9 in iPSC-derived neurons of FTD/ALS patients with TDP-43 mutations. *PLoS One* **8**, e76055 (2013).
146. Morlando, M. *et al.* FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *EMBO J.* **31**, 4502–4510 (2012).
147. Freischmidt, A., Müller, K., Ludolph, A. C. & Weishaupt, J. H. Systemic dysregulation of TDP-43 binding microRNAs in amyotrophic lateral sclerosis. *Acta Neuropathol Commun* **1**, 42 (2013).
148. Figueroa-Romero, C. *et al.* Expression of microRNAs in human post-mortem amyotrophic lateral sclerosis spinal cords provides insight into disease mechanisms. *Mol. Cell. Neurosci.* **71**, 34–45 (2016).
149. Butovsky, O. *et al.* Modulating inflammatory monocytes with a unique microRNA gene signature ameliorates murine ALS. *J. Clin. Invest.* **122**, 3063–3087 (2012).
150. Koval, E. D. *et al.* Method for widespread microRNA-155 inhibition prolongs survival in ALS-model mice. *Hum. Mol. Genet.* **22**, 4127–4135 (2013).
151. Sasaki, T., Shiohama, A., Minoshima, S. & Shimizu, N. Identification of eight members of

the Argonaute family in the human genome☆. *Genomics* **82**, 323–330 (2003).
152. Paez-Colasante, X., Figueroa-Romero, C., Sakowski, S. A., Goutman, S. A. & Feldman, E. L. Amyotrophic lateral sclerosis: mechanisms and therapeutics in the epigenomic era. *Nat. Rev. Neurol.* **11**, 266–279 (2015).
153. Koritzinsky, M. *et al.* Phosphorylation of eIF2alpha is required for mRNA translation inhibition and survival during moderate hypoxia. *Radiother. Oncol.* **83**, 353–361 (2007).
154. Spriggs, K. A., Bushell, M. & Willis, A. E. Translational regulation of gene expression during conditions of cell stress. *Mol. Cell* **40**, 228–237 (2010).
155. Kedersha, N. L., Gupta, M., Li, W., Miller, I. & Anderson, P. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J. Cell Biol.* **147**, 1431–1442 (1999).
156. Harding, H. P. *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* **6**, 1099–1108 (2000).
157. Mazroui, R. *et al.* Inhibition of ribosome recruitment induces stress granule formation independently of eukaryotic initiation factor 2alpha phosphorylation. *Mol. Biol. Cell* **17**, 4212–4219 (2006).
158. Colombrita, C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051–1061 (2009).
159. Dewey, C. M. *et al.* TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol. Cell. Biol.* **31**, 1098–1108 (2011).
160. Liu-Yesucevitz, L. *et al.* Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* **5**, e13250 (2010).
161. McDonald, K. K. *et al.* TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum. Mol. Genet.* **20**, 1400–1410 (2011).
162. Klann, E. & Dever, T. E. Biochemical mechanisms for translational regulation in synaptic plasticity. *Nat. Rev. Neurosci.* **5**, 931–942 (2004).
163. Mohr, E., Fehr, S. & Richter, D. Axonal transport of neuropeptide encoding mRNAs within the hypothalamo-hypophyseal tract of rats. *EMBO J.* **10**, 2419–2424 (1991).
164. Willis, D. *et al.* Differential transport and local translation of cytoskeletal, injury-response, and neurodegeneration protein mRNAs in axons. *J. Neurosci.* **25**, 778–791 (2005).
165. Ainger, K. *et al.* Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. *J. Cell Biol.* **123**, 431–441 (1993).
166. Huang, Y.-S., Carson, J. H., Barbarese, E. & Richter, J. D. Facilitation of dendritic mRNA transport by CPEB. *Genes Dev.* **17**, 638–653 (2003).
167. Krichevsky, A. M. & Kosik, K. S. Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation. *Neuron* **32**, 683–696 (2001).
168. Alami, N. H. *et al.* Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* **81**, 536–543 (2014).
169. Fallini, C., Bassell, G. J. & Rossoll, W. The ALS disease protein TDP-43 is actively transported in motor neuron axons and regulates axon outgrowth. *Hum. Mol. Genet.* **21**, 3703–3718 (2012).
170. Wang, I.-F., Wu, L.-S., Chang, H.-Y. & Shen, C.-K. J. TDP-43, the signature protein of FTLD-U, is a neuronal activity-responsive factor. *J. Neurochem.* **105**, 797–806 (2008).
171. Zhang, Y.-J. *et al.* Progranulin mediates caspase-dependent cleavage of TAR DNA binding

protein-43. *J. Neurosci.* **27**, 10530–10534 (2007).
172. Johnson, B. S. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* **284**, 20329–20339 (2009).
173. Bigio, E. H. *et al.* Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathologica* **125**, 463–465 (2013).
174. Mompeán, M., Baralle, M., Buratti, E. & Laurents, D. V. An Amyloid-Like Pathological Conformation of TDP-43 Is Stabilized by Hypercooperative Hydrogen Bonds. *Front. Mol. Neurosci.* **9**, 125 (2016).
175. Zacco, E., Martin, S. R., Thorogate, R. & Pastore, A. The RNA-Recognition Motifs of TAR DNA-Binding Protein 43 May Play a Role in the Aberrant Self-Assembly of the Protein. *Frontiers in Molecular Neuroscience* **11**, (2018).
176. Wang, Y.-T. *et al.* The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J. Biol. Chem.* **288**, 9049–9057 (2013).
177. Chang, C.-K. & Huang, T.-H. Untangling the structure of the TDP-43 N-terminal domain. *FEBS J.* **283**, 1239–1241 (2016).
178. Afroz, T. *et al.* Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature Communications* **8**, (2017).
179. Hergesheimer, R. C. *et al.* The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain* **142**, 1176–1194 (2019).
180. Igaz, L. M. *et al.* Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. *J. Clin. Invest.* **121**, 726–738 (2011).

Figures

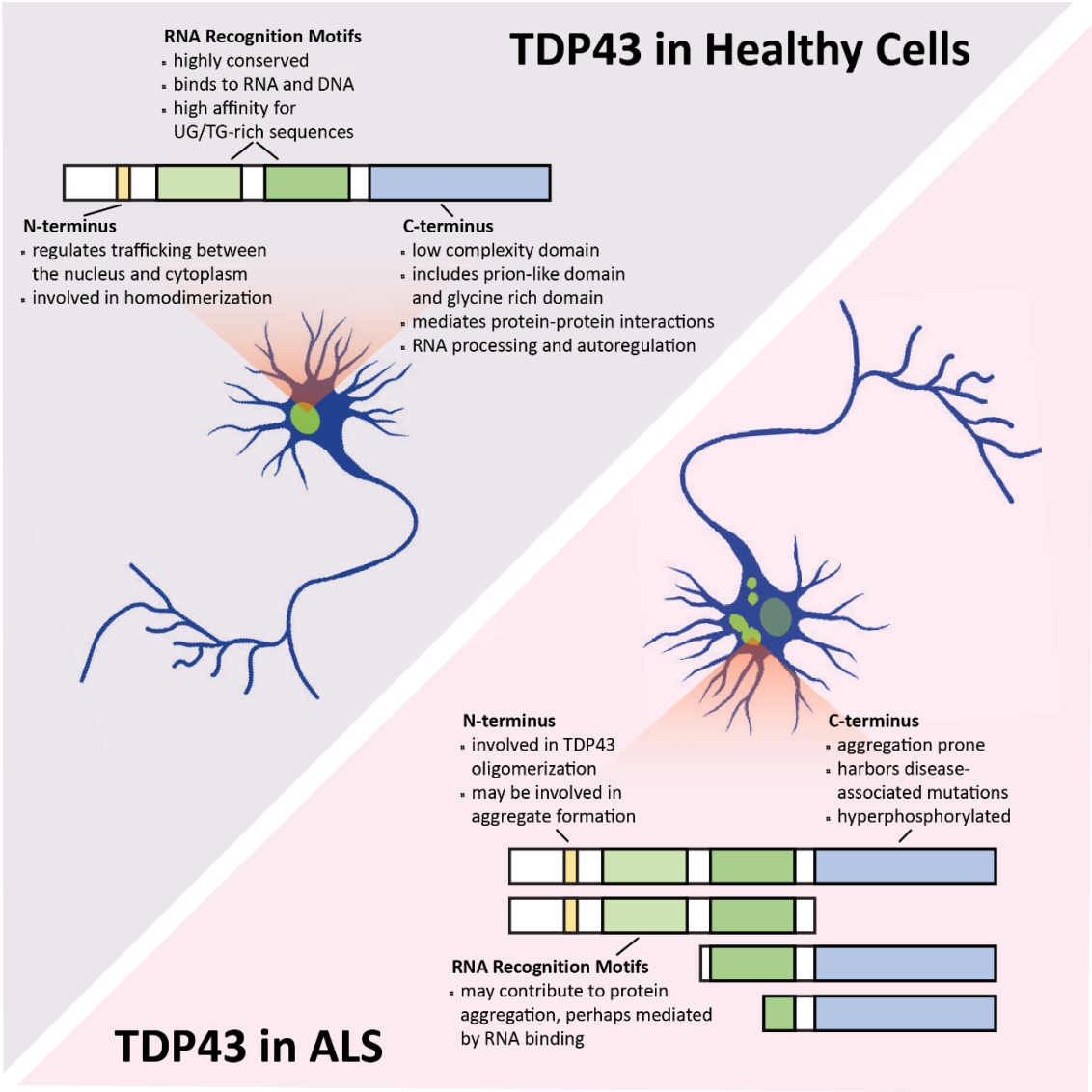


Figure 1.1. TDP43 structure and function. TDP43 consists of an N-terminus that contains a nuclear export sequence (yellow), two RNA-recognition motifs (green) and a C-terminus (blue) that encompasses both a prion-like domain and a glycine rich domain. In healthy cells (top left) TDP43 is primarily nuclear and plays critical roles in RNA processing and metabolism. However, in the majority of ALS cases TDP43 is mislocalized from the nucleus to the cytoplasm, where it forms cytoplasmic aggregates (bottom right). These aggregates contain several TDP43 species, and may contribute to ALS pathogenesis through a gain of toxic TDP43 function, a loss of normal TDP43 function, or a combination of the two.

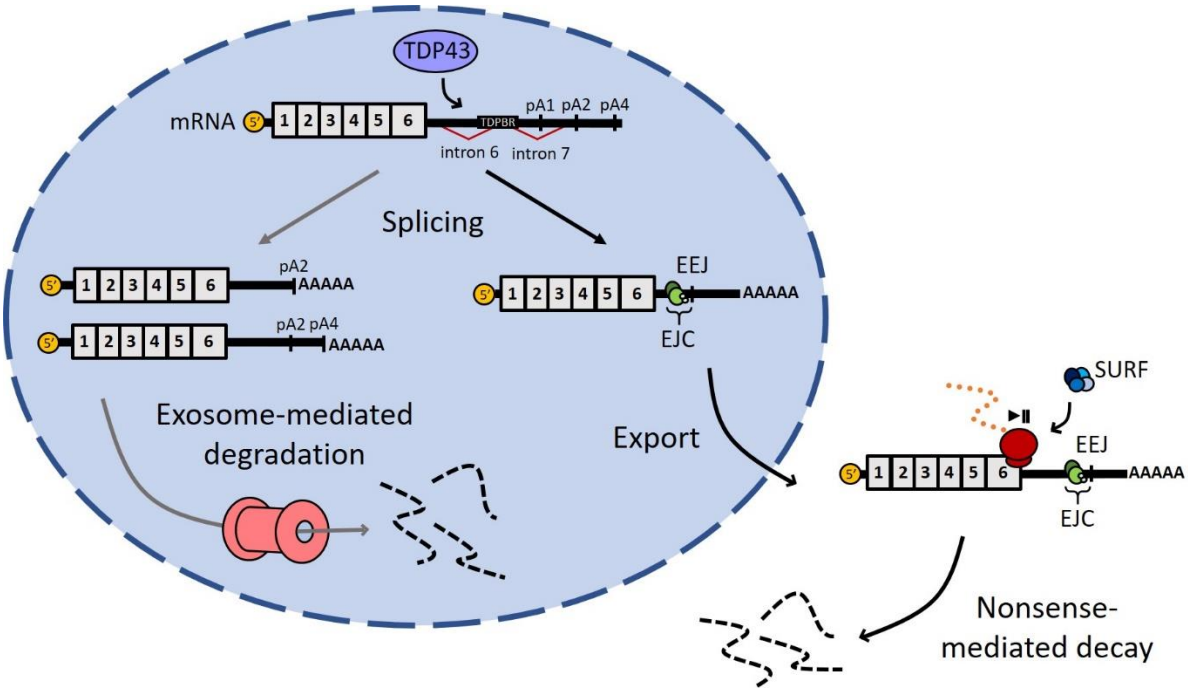


Figure 1.2. TDP43 autoregulation. TDP43 may destabilize its own mRNA transcript through two distinct mechanisms. In the first (gray arrows), TDP43 protein recognizes the TDP43 binding region (TDPBR) within the 3' UTR of its own transcript, stimulating the removal of alternative intron 7 and the primary polyadenylation site (pA1) contained within the intron. Spliced transcripts are preferentially retained in the nucleus and targeted for exosome-mediated decay. In the second mechanism (black arrows), the removal of introns 6 and/or 7 creates exon-exon junctions (EEJs) and the assembly of exon junction complexes (EJCs). The transcript is then exported to the cytoplasm. During the first or pioneer round of translation, the ribosome pauses at the stop codon, allowing the association of the SURF complex with the ribosome. Factors within the downstream EJC interact with UPF1 in the SURF complex, triggering UPF1 phosphorylation and nonsense-mediated mRNA decay.

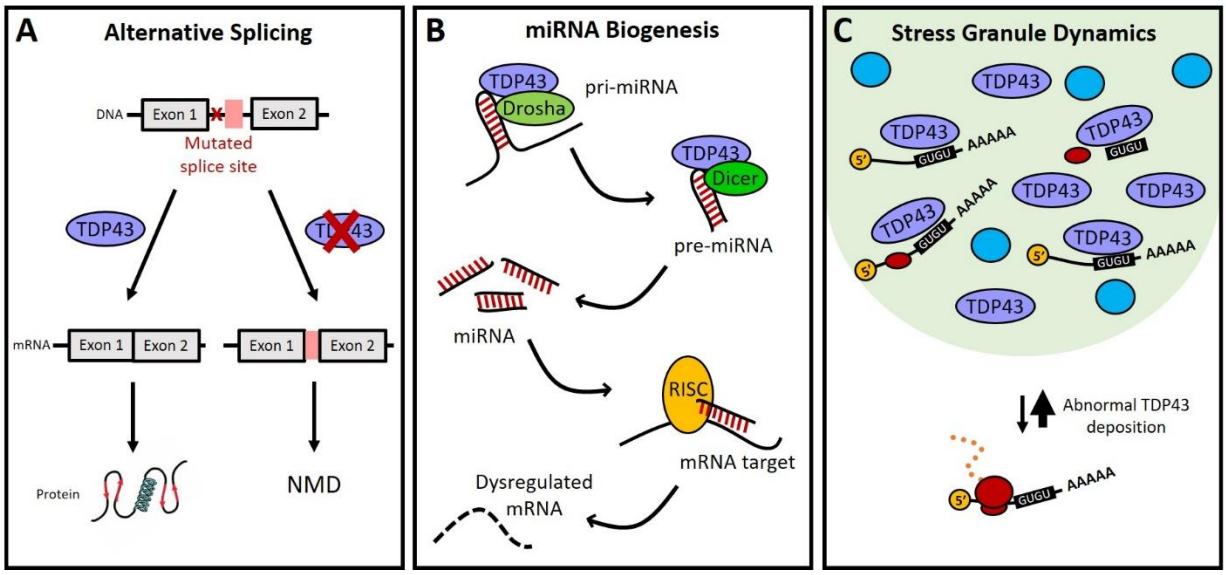


Figure 1.3. TDP43 deposition impacts RNA stability through several pathways. (A) Alternative splicing. Mutations that introduce novel splice sites can lead to the inclusion of unannotated or "cryptic" exons (pink box). These faulty transcripts are often targeted by NMD. Typically, TDP43 is a strong repressor of these unannotated splicing events, but nuclear exclusion prevents TDP43 from performing this function, and abnormal transcripts accumulate. (B) miRNA biogenesis. TDP43 promotes several steps of miRNA biogenesis, and regulates the formation of key miRNAs that, in turn, control the stability and translation of mRNAs that are essential for neuronal survival, growth and development. (C) Stress granule dynamics. TDP43 is one of several RNA-binding proteins (blue circles) that localize to SGs (light green) in response to various conditions. Because TDP43 recognizes thousands of GU-rich transcripts, cytoplasmic TDP43 deposition within SGs forces mRNA recruitment to SGs, shifting transcripts from actively translating polysomes to inert, though stable, SGs.

Chapter 2. RNA Degradation in Neurodegenerative Disease*

2.1 Introduction

Ribonucleic acid (RNA) homeostasis is dynamically modulated in response to changing physiological conditions. Tight regulation of RNA abundance through both transcription and degradation determines the amount, timing, and location of protein translation. This balance is of particular importance in neurons, which are among the most metabolically active and morphologically complex cells in the body. As a result, any disruptions in RNA degradation can have dramatic consequences for neuronal health. In this chapter, we will first discuss mechanisms of RNA stabilization and decay. We will then explore how the disruption of these pathways can lead to neurodegenerative disease.

2.2 Mechanisms to Maintain RNA Stability

Following transcription, the newly formed transcript can be stabilized in several ways (Figure 2.1). Most RNA that codes for protein, also referred to as coding or messenger RNA (mRNA), undergoes several processing steps that prevent degradation, assist in export from the nucleus, and aid in translation. Additionally, both coding and non-coding RNA (ncRNA) are

* This chapter represents the following manuscript:

Weskamp K, Barmada SJ. (2018). RNA Degradation in Neurodegenerative Disease. *Advances in neurobiology*. https://doi.org/10.1007/978-3-319-89689-2_5

stabilized by the adoption of unique secondary structures or sequestration in cytoplasmic ribonucleoprotein particles when the cell is under stress.

2.2.1 Polyadenylation

Polyadenylation refers to the addition of a series of adenosine monophosphates to the 3' end of mRNA transcripts[1]. This poly(A) tail protects nascent mRNA from enzymatic degradation[2,3], facilitates nuclear export[4], and assists in translation[3]. Polyadenylation begins when a complex of several proteins recognizes a binding site on the mRNA transcript. An enzyme in this complex, cleavage/polyadenylation specificity factor (CPSF), cleaves the 3' end of the transcript, and a second component, polyadenylate polymerase, adds sequential adenosine monophosphate units to create the poly(A) tail[5]. As the poly(A) tail grows longer, polyadenylate binding protein 2 (PAB2) is recruited, which further increases the affinity of polyadenylate polymerase to the RNA[6]. Additional poly(A)-binding proteins then associate with the tail and facilitate nuclear export, stabilization of the RNA, and translation[7].

Many transcripts harbor more than one polyadenylation site. The site that is ultimately utilized primarily affects the length of the 3' untranslated region (UTR), with little direct influence on protein translation or function[8]. However, the 3'UTR may also encode microRNA recognition elements[9], DNA methylation sites[10], or motifs recognized by regulatory RNA-binding proteins[11,12]. Thus, where a poly(A) tail starts can significantly influence the likelihood of transcript degradation. Moreover, in some cases alternative poly(A) binding sites occur within the coding region, and their usage results in truncation of the translated protein[13]. Poly(A) tails are gradually eroded over time, and transcripts with shorter tails are both less likely to be transcribed

and more likely to be degraded[14]. This process can be accelerated by the binding of microRNA to the 3' UTR or through the removal or degradation of poly(A) binding proteins[15].

2.2.2 Methylguanine Cap

The majority of coding RNAs undergo a second processing step that involves the addition of a methylguanine cap to the 5' end of the transcript. This cap stabilizes the transcript by preventing exonuclease-mediated degradation[16–18], and is also required for the translation of most mRNAs[19,20]. Additionally, the 5' cap assists in splicing[21–25], nuclear export[24,25], and possibly polyadenylation[26].

The capping process is initiated before transcription is complete, and begins when RNA triphosphatase removes one of the 5' terminal phosphate groups[27]. mRNA guanylyltransferase then catalyzes the addition of guanosine triphosphate to the remaining terminal biphosphate to create an unusual 5' to 5' triphosphate linkage. This guanosine is then methylated by a methyltransferase[27]. The cap binding complex (CBC) binds to the methylated 5' cap, which is in turn recognized by the nuclear pore complex and exported into the cytoplasm[28,29]. Once there, the CBC is replaced by the translation factors eIF4E and eIF4F, which are recognized by other translation initiation machinery components, including the ribosome[30,31].

Binding of the CBC and translation factors also stabilize transcripts by blocking the binding of decapping enzymes[32–34]. When these decapping enzymes outcompete the translation factors, they hydrolyze the 5' cap and expose the 5' monophosphate. The resulting decapped transcripts are subject to rapid degradation by 5' exonucleases[35].

2.2.3 Secondary Structure

DNA primarily forms double helices, but the single stranded nature of RNA and its propensity to form hydrogen bonds allows it to form more complex structures that can directly affect transcript stability. The most common RNA secondary structure is the hairpin loop, created when two complementary regions of the same strand base-pair to form a double helix that ends in an unpaired loop[36]. These loops are found in pre-microRNA, transfer RNA (tRNA), and mRNA, and their stability depends on several factors, including length, degree of complementarity in the stem, and guanine to cytosine base pair content. Hairpin loops stabilize mRNA[37–40] and in many cases increase translation efficiency[39,40]. This may occur by blocking exonuclease activity, but the precise mechanism remains unclear. Hairpin loops may also act as binding sites for proteins that direct mRNA transport and localization[41–43].

The combination of several hairpin loops forms a multiloop; the most abundant example of this structure is found in the cloverleaf-shaped tRNAs that assist in protein translation. The relative stabilities of multiloops vary based on size, number of loops, and complementarity[44]. Hairpin loops can also form pseudoknots, in which at least two hairpin loops are linked by single stranded loops. Pseudoknots are relatively stable, and though little is known about their functional significance, they form the catalytic core of some ribozymes[45,46] and telomerases[47] and may also be involved in translation[48]. Other structures, such as G-quadruplexes and R-loops, are more often associated with disease and will be discussed below.

2.2.4 Stress Granules

Cells undergo a wide range of molecular changes in response to environmental stressors, including the inhibition of conventional translation[49,50] and the formation of stress granules (SGs),

cytoplasmic ribonucleoprotein particles rich in mRNA, RNA-binding proteins, and stalled translation initiation complexes[51–53]. SG coalescence effectively sequesters the attached mRNAs and the 40S ribosome subunit[54,55], preventing further translation and stabilizing the bound mRNAs. Proteins unrelated to the original translation initiation complex are also recruited, and their composition helps determine SG dynamics and longevity[56]. Which proteins participate is often dependent on their posttranslational modifications and the specific stressor involved[57–61], providing a rapid and reversible way for the cell to modulate SG formation and composition. Many RNA-binding proteins found in SGs contain low-complexity domains that are inherently flexible; the ability of these domains to form reversible homo- and heterotypic interactions with one another via their low-complexity domains may be responsible for the dynamics of SG formation and dissociation[62,63]. Additionally, SGs often contain a number proteins that promote RNA stability and regulate translation[64]. Moreover, deadenylation is largely inhibited in stress granules[65–67]. When the stressor has passed, several RNA-binding proteins catalyze SG disassembly[68–70], and the transcript is either degraded or released to resume translation. These observations suggest that SGs serve two basic functions: preventing the translation of unnecessary transcripts during stress, and protecting these transcripts from degradation until the stress has subsided.

2.3 Mechanisms of RNA Decay

The typical life of an mRNA transcript includes a complex sequence of events including transcription, capping, adenylation, splicing, and export. When mistakes occur during this process, quality control mechanisms exist to recognize and eliminate defective transcripts that may give rise to dysfunctional or toxic proteins (Figure 2.1). However, these pathways do more than ensure the fidelity of RNA transcripts. They also serve important regulatory roles, enabling rapid

modulation of steady-state RNA levels—and therefore protein production—in response to changes in the intracellular or extracellular environment.

2.3.1 RNA Degradation Machinery

There are three major classes of intracellular RNA-degrading enzymes: endonucleases that cut RNA internally, 5' to 3' exonucleases that degrade RNA from the 5' end, and 3' to 5' exonucleases that hydrolyze RNA from the 3' end. These enzymes may work independently or within a complex such as the exosome, a versatile structure for the degradation of immature or abnormal RNA. The core of the eukaryotic exosome complex is formed by nine proteins, six of which are members of the RNase PH-like family[71]. These form a ring that is capped by three additional proteins with RNA-binding domains[72]; this structure bears remarkable similarity to the 26S proteasome[73], which consists of a central proteolytic barrel (the 20S core) capped on either end by 19S regulatory subunits. The exosome is primarily composed of 3'-5' exoribonucleases, and RNAs are degraded by removing terminal nucleotides from the 3' end of the transcript. This occurs through the cleavage of phosphodiester bonds, either through RNase PH-like protein-mediated phosphorolytic cleavage or hydrolytic cleavage by proteins associated with the exosome[74]. Several other proteins bind to the exosome to regulate its activity and specificity[75–77]. The exosome also processes small nuclear RNAs, small nucleolar RNAs, and ribosomal RNAs[78], though how these molecules are targeted to and released from the exosome remains unclear.

2.3.2 Nonsense-mediated Decay

Occasionally, errors introduced during transcription, insertions, deletions or nonsense mutations uncover premature stop codons (PTCs) within the coding sequence of an mRNA. If translated, PTC-containing transcripts would encode truncated proteins that may have toxic gain-

of-function or dominant-negative activities. Nonsense-mediated decay (NMD) is a surveillance mechanism that eliminates transcripts containing PTCs, thereby preventing the synthesis of proteins that could be detrimental to the cell.

mRNA transcripts undergo splicing following transcription, during which introns are removed and exons are spliced together. The resulting exon exon junctions (EEJs) are occupied by a complex of proteins (the exon junction complex, or EJC) that assist in splicing until they are displaced by the ribosome during the first, or pioneer, round of translation. If the stop codon is downstream or within about 50 nucleotides of the final EJC, the transcript is translated normally. According to the EJC model of NMD, a stop codon that occurs upstream of an EJC is recognized as a PTC, triggering transcript degradation[79,80]. When the ribosome stalls at a PTC, the protein UPF1, along with the eukaryotic release factors eRF1 and eRF3, form the surveillance complex (SURF) and bind adjacent to the PTC. SURF then interacts with two components of the nearby EJC, UPF2 and UPF3B[81–83]. This triggers UPF1 phosphorylation, which causes the complex to move along the mRNA, resolving secondary structure and removing adherent proteins that may inhibit degradation[84,85]. Phosphorylated UPF1 also binds to SMG6, an endonuclease that directly cleaves the mRNA[86,87], as well as SMG5 and SMG7, which trigger deadenylation[88], decapping, and further degradation[89]. Additionally, UPF1 may be recruited to transcripts independent of a PTC or adjacent EJC, particularly within long 3' UTRs[90]. A working theory is that UPF1 preferentially binds long 3' UTRs and is phosphorylated via an unknown mechanism, triggering transcript decay. However, more work is required to identify the pathway resulting in destabilization of transcripts bearing long 3'UTRs.

2.3.2.1 Alternative Exon Inclusion and Exclusion

Though NMD is an important quality control mechanism, it also helps regulate the expression of functional mRNA[91], predominantly through alternative mRNA splicing. This phenomenon is remarkably widespread: NMD-related regulation of transcript abundance is involved in cell proliferation[92,93], immunity[94], stress[95], viral response[96], and neuronal activity[97,98]. The differential inclusion or exclusion of exons (alternative splicing) enables a single gene to encode multiple transcript and protein isoforms, and in many cases alternatively spliced transcripts are subject to NMD. Because changes in the splicing environment determine which isoforms are produced[99,100], alternative splicing can regulate gene expression by creating transcripts that are more or less stable. An estimated 33% of alternative transcripts contain PTCs[101], and between 12% and 45% of alternatively spliced transcripts are estimated to be NMD targets[101]. Regulated unproductive splicing (RUST) of this type regulates RNA abundance in relation to neuronal activity levels[102], developmental stage, and cell type[103]. Moreover, there is growing evidence that RUST is utilized by several RNA-binding proteins to regulate their own expression (autoregulation), particularly components of the splicing machinery[104–108].

2.3.2.2 Upstream Open Reading Frames

Upstream open reading frames (uORFs) are mRNA elements that include a start codon in the 5’ UTR that is out-of-frame with the main coding sequence. Because ribosomes bind to the 5’ cap of the mRNA and scan for start codons, uORFs can disrupt or interfere with translation of the downstream coding sequence[109,110]. Moreover, any stop codon at the 3’ end of the uORF resembles a PTC within the context of the whole transcript. As predicted by the EJC model of NMD, the

presence of uORFs correlate with lower expression levels of the downstream ORF[111,112], and uORF-bearing transcripts are particularly susceptible to degradation by NMD[113–115].

2.3.3 Nonstop Decay

Nonstop decay (NSD) is a surveillance mechanism involved in the detection and degradation of mRNA transcripts that lack stop codons[77,116] due to premature polyadenylation or point mutations that disrupt existing terminal codons. Without a recognizable stop codon, the ribosome translates into the poly(A) tail and then stalls, unable to release the mRNA transcript[117].

NSD is activated when Ski7, a component of the exosome complex, binds the empty aminoacyl (A) site of the stalled ribosome via its C-terminal domain[76,77]. This is supported by the fact that C-terminal deletions of Ski7 result in impaired NSD but do not affect general exosome function[116]. Additionally, the Ski7 C-terminal domain strongly resembles other proteins that bind the ribosome during normal translation, elongation, and termination such as EF1a and eRF3[118]. After binding, Ski7 releases the stalled ribosome and recruits the exosome to rapidly deadenylate the transcript[77,116,119,120].

2.3.4 No-go Decay

No-go decay (NGD) is a mechanism that recognizes mRNA transcripts stalled during translation[121–123] due to damaged RNA, stress[124], or strong secondary structure that blocks the progress of translation machinery[121]. NGD is the most recently discovered RNA surveillance pathway, and as such little is known about its mechanism. However, evidence suggests that NGD may degrade mRNA in a manner that resembles translation termination. Two proteins that promote

NGD, Hbs1 and Dom34, strongly resemble eRF1 and eRF3, two factors that catalyze the end of translation[121,125].

Analogous to Ski7 in NSD, Hbs1 possesses the same C-terminal domain that allows EF1a, eRF3, and Ski7 to bind the empty A site on the stalled ribosome[126,127]. Dom34 is homologous to eRF1 and binds directly to Hbs1[126,128]. Upon binding, the Dom34/Hbs1 complex triggers the release of the nascent peptide and the ribosome is released or degraded. Likewise, the mRNA transcript is targeted for endonucleolytic cleavage and the fragments are subsequently degraded via the exosome or exonucleases[121,125]. It is not currently known how the Dom34/Hbs1 complex releases the mRNA from the ribosome, but the close relation between Hbs1 and Ski7 suggests that ribosome release may occur in the same manner as NSD. Moreover, NGD can occur independently of the Dom34/Hbs1 complex; further work is needed to identify the other factors involved.

Additionally, it remains unclear why some transcripts are targeted by NGD and not others. Pausing during translation is a normal occurrence[129] and may even serve biological functions[130–132], but only a fraction of transcripts are NGD substrates. Potentially important factors include the degree of ribosome stalling and whether or not the A site is empty to allow Dom34/Hbs1 complex binding. Further studies are needed to clarify this mechanism.

2.3.5 Adenylate-Uridylate-Rich Elements

While some mRNA decay pathways target faulty transcripts, others allow the cell to rapidly modulate gene expression in response to intracellular and extracellular stimuli. Several of these pathways regulate transcript levels via binding sites within the 3’ UTR, including adenylate-uridylate-rich elements (AREs), Staufen-mediated decay, microRNAs, and constitutive decay elements.

AREs are 50-150 nucleotide regions with frequent adenine and uridine bases that generally target the mRNA for rapid degradation[133,134]. The mechanism underlying this pathway is not well understood, but several RNA-binding proteins interact with these sites and modulate transcript stability. For example, overexpression of hnRNP D, also known as ARE RNA binding protein 1 (AUF1), destabilizes mRNA containing AREs[135,136]. Conversely, AUF1 depletion increases both ARE-containing mRNA stability and abundance of the corresponding proteins[137,138]. Similarly, ablation of tristetraprolin (TTP), an RNA-binding protein that also recognizes AREs, increases mRNA and protein levels in a variety of cell types[139–141] and transcripts[142–147].

Though the exact mechanism is unclear, the association of ARE-binding proteins to AREs is followed by deadenylation[148–151], decapping, and 3' to 5' degradation via the exosome[152]. Certain subunits of the exosome bind to AREs directly, and several ARE-binding proteins including TTP associate with the exosome *in vitro*[75,153], ensuring rapid and preferential elimination of ARE-continuing transcripts. Many ARE-binding proteins are also associated with SGs and P-bodies (discussed later in this chapter), suggesting that 5' to 3' exonuclease-mediated degradation may contribute to the turnover of ARE-containing transcripts as well[154,155]. However, not all ARE-binding proteins trigger mRNA decay. For example, the Hu family of proteins stabilize bound ARE-containing transcripts[156–159], suggesting that the effect of AREs on RNA stability depends on a combination of factors, including the ARE-binding protein, transcript, and environment.

2.3.6 Staufen-mediated Decay

Staufen-mediated decay (SMD) also regulates transcript levels via the 3' UTR. SMD is triggered when Staufen-1 (Stau1) recognizes double stranded RNA structures that form sufficiently downstream of the termination codon[160,161]. Staufen binding sites (SBS) are created by

intramolecular hairpin loop formation within the 3' UTR[161], or intermolecular base-pairing of the 3' UTR with partially complementary long noncoding RNA[162]. Upon binding to the SBS, Stau1 recruits UPF1, which in turn stimulates mRNA decay[160], likely in much the same way as in NMD. Moreover, given that UPF1 is critical for both SMD and NMD, there may be competition between the two pathways based on the availability of UPF1[163].

2.3.7 microRNAs

microRNAs (miRNAs) are small, non-coding RNAs that base-pair with complementary sequences within RNA transcripts to trigger their decay and/or translational repression. These 20-25 nt RNAs are produced from an RNA precursor (pri-miRNA) that forms a hairpin loop shortly after transcription[164,165]. This structure is recognized by the nuclear protein DGCR8, which recruits the enzyme Drosha to cleave the hairpin from the rest of the transcript[166,167]. The resulting molecule (pre-miRNA) is then exported to the cytoplasm[168] where the enzyme DICER cuts away the looped end[169], leaving a duplex of two short, complementary RNA strands behind. Though either strand can function as a mature miRNA, one is usually degraded[170,171]. The remaining miRNA associates with the RNA-induced silencing complex (RISC), which assists in orienting the miRNA with its mRNA target, repressing translation of the target transcript and triggering its degradation.

The bound miRNA guides RISC to its binding site (miRNA recognition element or MRE) on the target transcript, most often within the 3' UTR, though binding can occur within coding regions as well[172,173]. The degree of miRNA-mRNA complementarity is a major predictor of transcript fate[174]. High degrees of sequence complementarity allow the Argonaute family of proteins—components of RISC[175]—to catalyze RNA decay through an unknown mechanism that may involve deadenylation, decapping, or exonucleolytic degradation[176,177]. In contrast, miRNAs

that bind weakly or with less complementarity induce translational repression[174] through a mechanism that remains unclear.

2.3.8 Constitutive Decay Elements

In addition to AREs, SBSs, and MREs, structured RNA degradation motifs may directly lead to transcript turnover. Constitutive decay elements (CDEs) are stem loop structures located within the 3' UTR that trigger mRNA decay[178,179] through recruitment of the RNA-binding protein Roquin1[179,180]. Roquin1 binds to the CDE stem loop structure via two binding sites in its ROQ domain[180], triggering degradation by recruiting the Ccr4-Caf1-Not deadenylation complex[179]. A transcriptome-wide search of 3' UTRs in mice revealed several unique CDEs that are frequent and highly conserved across vertebrate species. Many, but not all, of these CDEs are Roquin1-associated[179], indicative of potential novel and unexplored pathways responsible for RNA decay.

2.3.9 Histone mRNAs

Much like CDE-containing transcripts, histone mRNAs encode highly conserved stem loop structures within their 3' UTRs. These hairpins are essential for the rapid synthesis and degradation of histone mRNA during the S phase of the cell cycle, during which the cell undergoes DNA replication and chromosome remodeling[181]. At the end of S phase, histone hairpin loops are recognized by stem loop binding protein (SLBP), which recruits the proteins necessary to add a short, oligonucleotide tail to histone mRNAs[182]. The oligonucleotide tail forms a binding site for LSM1-7, which triggers degradation via the exosome and endonucleases[182]. Interestingly, histone mRNA decay also requires UPF1 and its interaction with SLBP[183], though the exact role of UPF1 in histone mRNA metabolism remains unclear.

2.3.10 Processing Bodies

Processing bodies (P-bodies) are dynamic cytoplasmic foci comprised of mRNA and RNA-binding proteins. While SGs primarily sequester and protect mRNA until it can resume translation, P-bodies target associated transcripts for translational repression, decapping, and decay. Although P-body assembly is not required for RNA decay[184], it may directly compete with translation initiation; only transcripts that are not engaged in translation can be recruited to P-bodies[185–187], and upon translational inhibition P-bodies increase in number[185,188]. Conversely, a decrease in P-body components leads to an increase in mRNAs associated with actively-translating polysomes[189]. P-bodies lack translation initiation machinery[185,187], and are instead primarily composed of proteins associated with translational repression and mRNA decay, including decapping enzymes, exonucleases and NMD components[190]. This suggests that functional transcripts undergo active translation before they are recruited to P-bodies. Once transferred, the mRNA is no longer translated[189,191] and is instead degraded by decapping enzymes[192,193] or other nucleases. However, mRNAs may also escape P-bodies and resume translation[187,194], and regulated expression of proteins such as NoBody and MLN51 can drive P-body disassembly[195,196]. Together, these observations indicate that P-bodies are part of a highly dynamic process characterized by constant flux between pools of mRNA transcripts that are being actively translated, those that are stalled or sequestered in SGs, and those that are being degraded within P-bodies.

2.4 RNA Turnover in Neurodegenerative Disease

The regulation of RNA is critical to cell health, and increasing evidence indicates that disruption of RNA stability may underlie neurodegenerative disease. Alterations in RNA turnover

have been identified in several pathways, including RNA sequestration in stress granules or foci, RNA transport, the exosome, alternative splicing, and retrotransposons (Figure 2.2).

2.4.1 RNA Sequestration

During times of stress, the cell diverts its energy and resources towards survival and recovery. A powerful mechanism to conserve resources is the sequestration of mRNAs in SGs to limit the translation of nonessential proteins. Typically, when the stressor passes, SGs dissolve and stalled mRNAs are released for translation. However, during prolonged periods of stress or disease, SGs sometimes fail to disassemble. This extended sequestration of mRNAs could effectively disrupt the delicate balance between SGs, polysomes, and P-bodies, effectively interrupting mRNA homeostasis, interfering with protein synthesis and potentially contributing to downstream toxicity in neurodegenerative diseases.

2.4.1.1 Disruption of Stress Granule Dynamics

Of the ~125 proteins identified as components of human SGs, 60% are RNA-binding proteins[197]. This group of proteins is also highly enriched for the low complexity domains that facilitate the reversible aggregation of proteins into membraneless organelles such as SGs. The mutation or mislocalization of several RNA-binding proteins stabilize SGs, sometimes driving them to form irreversible aggregates that sequester mRNA and RNA-binding proteins indefinitely and disrupt SG homeostasis. Conversely, though the machinery that drives SG disassembly remains unclear, any errors within this pathway may likewise lead to RNA dyshomeostasis and subsequent disease.

2.4.1.2 RNA-binding Proteins in Stress Granule Dynamics

TDP43 and FUS are two stress granule components that are integrally involved in neurodegenerative disease, particularly amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Both TDP43 and FUS are primarily nuclear proteins, but their cytoplasmic mislocalization[198–200] and nuclear exclusion[201–203] are characteristic features of ALS and FTD. These proteins are capable of nucleocytoplasmic shuttling: in response to various stressors they associate with cytoplasmic SGs, but when the stress has passed they return to the nucleus[204]. ALS-linked mutations in the genes encoding TDP43 and FUS promote increased association with SGs[202,205], abnormal SG formation[206], and reduced SG dissociation[207,208]. TDP43 and FUS play important roles in alternative splicing and the stress response, and their sequestration impacts the processing of several transcripts that are critical for neuronal viability[209,210]. Likewise, excess cytoplasmic TDP43 and FUS may sequester related RNA-binding proteins within SGs, further disrupting RNA homeostasis[64]. Importantly, TDP43- and FUS-related toxicity relies upon the ability of these proteins to bind RNA. Deletion of the RNA recognition-motifs in either protein greatly reduces toxicity without affecting localization[211,212], suggesting that RNA binding, not localization, imparts toxicity. Furthermore, these observations indicate that the sequestration of mRNAs themselves, not just RNA-binding proteins, is particularly damaging to neurons.

ALS-linked mutations are also found in other RNA-binding proteins such as Matrin3[213], hnRNPA1, hnRNPA2/B1[214], and TIA1[215], all of which associate with SGs. These mutations are often centralized within the proteins' low complexity domains, and evidence indicates that they likewise alter SG dynamics, suggesting a link between SG association/dissociation and pathogenicity.

2.4.1.3 Stress Granule Disassembly

Though relatively little is known about SG disassembly, evidence suggests that valosin-containing protein (VCP) is crucial for this phenomenon. VCP regulates several cellular processes including autophagy[216], chromatin remodeling[217], and membrane trafficking[216], as well as SG clearance[218]. VCP accumulates in SGs, and its knockdown results in the persistence of SGs even after the stressor has passed[218]. Moreover, mutations in the gene encoding VCP cause a multisystem proteinopathy that includes ALS and FTD[219], and the overexpression of mutant VCP results in impaired SG disassembly[218]. Thus, pathogenic mutations in the genes encoding VCP, TDP43, and FUS all stabilize SGs, thereby effectively sequestering essential mRNA and RNA-binding proteins within these organelles. A such, altered SG dynamics and abnormal RNA stability may represent a conserved pathway underlying ALS, FTD and related neurodegenerative diseases.

2.4.2 Nucleotide Repeats and RNA Foci

Microsatellites are repeated tracts of nucleic acids that compose approximately 50% of the human genome[220]. These regions are a source of genomic instability, and expansion mutations that increase the number of repeats above a certain threshold can lead to neurodegenerative diseases such as Huntington's disease (HD), myotonic dystrophy (DM), spinocerebellar ataxias, Freidrich's ataxia, fragile X syndrome, fragile X-associated tremor ataxia syndrome (FXTAS), ALS and FTD[221,222]. In most cases, the length of the expanded region is inversely correlated with prognosis — higher repeat number results in earlier onset and more severe symptoms. Repeat expansions have unique pathological implications— they form unique secondary structures that may disrupt translation, sequester RNAs and other proteins into nuclear foci, and serve as a substrate for non-canonical translation.

2.4.2.1 Repeat Expansion Secondary Structure

The majority of expansion mutations associated with disease are trinucleotide CNG repeats, where N is any nucleotide. Due to the high degree of complementarity, CCG, CAG, CUG, and CGG repeats readily form mismatched hairpin loops[223] whose stability increase proportionally with the number of repeats[224]. Tetra-, penta-, and hexa-nucleotide repeats also form hairpins[225], though they appear to be less stable.

Repeat expansions with a high percentage of guanine nucleotides can also form G-quadruplexes. In these structures, four guanine bases associate through Hoogsteen hydrogen bonding to form a square guanine tetrad, and two or more tetrads stack to form a G-quadruplex[226]. Whether or not G-quadruplexes exhibit a physiological function remains unknown, but some evidence indicates that they participate in transcriptional regulation and/or telomere maintenance[227]. They are also observed in association with cancer, copy number variants, and age-related disease, specifically ALS and FTD. The most common mutation responsible for inherited ALS and FTD consists of a GGGGCC (G_4C_2) repeat expansion in the first intron of *C9orf72*[228,229]. Unaffected individuals have 2-8 (G_4C_2) repeats[230], but tracts of > 32 (G_4C_2) repeats lead to ALS, FTD, or both with nearly 100% penetrance by age 80[231]. These repeats form stable G-quadruplexes[232], which are further stabilized in longer repeat expansions[233].

(G_4C_2) repeat expansions also form structures known as R-loops at the site of transcription, composed of nascently-synthesized RNA hybridized to the complementary DNA strand[234,235]. The unbound DNA strand may also form hairpins or G-quadruplexes, further stabilizing the loop[236]. In addition to *C9orf72*-related ALS/FTD, R-loops are also observed in fragile X syndrome and Freidrich's ataxia[237] characterized by CGG and GAA trinucleotide repeats, respectively. The

abundance of R-loops in these disorders depends on the size of the repeat expansion, with higher repeat number correlating with more frequent R-loops. These structures may contribute to the pathology of expansion diseases in several ways: by blocking translation[238], disrupting chromatin remodeling[239], or promoting genomic instability at the repeat expansion site[235]. In support of the pathogenic effects of R-loops, mutations in the gene encoding senataxin (*SETX),* a helicase that helps resolve R-loops[240], cause juvenile ALS (ALS4), while SETX overexpression prevents neurodegeneration in ALS models[241].

2.4.2.2 RNA Foci

In addition to their effects on RNA stability and translation, the propensity of repeat expansions to form stable secondary structures contributes to the formation of RNA foci[242,243]. These nuclear inclusions may drive pathogenesis through the sequestration and nuclear retention of specific RNA-binding proteins. For example, CUG repeat expansions in *DMPK* cause myotonic dystrophy type 1 (DM1), a neuromuscular disease characterized by progressive muscle loss and weakness. This repeat expansion sequesters and disrupts the splicing activity of muscleblind (MBNL)[244,245], a protein responsible for the processing of several key downstream transcripts[246]. MBNL binds to hairpins that result from repeat expansion mutations in *DMPK* with high affinity[245,247], and preventing MBNL sequestration via small molecules that recognize CUG hairpin loops restores its splicing activity and helps maintain RNA homeostasis in DM1 models[248]. Additionally, the RNA foci observed in DM1[249] and myotonic dystrophy type 2 (DM2)[250] sequester several other RNA-binding proteins, suggesting that global disruption of alternative splicing may contribute to DM pathogenesis[251]. RNA foci are also observed in *C9orf72*-linked ALS/FTD[252], where the G_4C_2 repeat transcripts sequester several splicing factors including

hnRNPA1, hnRNPH, and SC35, as well as the RNA-binding protein hnRNPA3 and the mRNA export receptor ALYREF[253]. The sequestration of proteins essential to multiple cellular processes by repeat expansion transcripts suggests that these diseases occur, at least in part, through an RNA gain of function mechanism.

2.4.2.3 Repeat Associated Non-AUG (RAN) Translation

Nucleotide repeats can be translated into polypeptides even if they are not located within a traditional open reading frame, via a non-canonical pathway termed repeat associated non-AUG (RAN) translation. RAN translation maybe triggered by hairpin loops formed by repeat-containing stretches of DNA, which effectively stall ribosome scanning and facilitate translational initiation at near-AUG codons[254–256]. This process occurs in multiple reading frames in both the sense and antisense directions, producing several dipeptide repeat-containing proteins (DPRs)[254]. RAN translation products are detected in spinocerebellar ataxia type 8, HD[257], DM1[254], FXTAS[256], and *C9orf72*-associated ALS/FTD[258], suggesting that RAN translation is a common phenomenon in repeat expansion diseases. In some cases, there appears to be an inverse relationship between RAN translation and RNA foci formed by repeat expansions. This observation suggests that the repeat-expanded RNA may be sequestered in nuclear foci, precluding nuclear export and subsequent translation[259]. This may serve as a coping response to prevent the translation of DPRs; failure of this coping response over time may result in increased RAN translation and subsequent neurodegeneration[260,261]. In support of this hypothesis, RNA foci in *C9orf72* mutant mice are abundant yet rarely associated with neurodegeneration[261]. RAN peptides may also affect RNA stability by disrupting membraneless organelles such as the nucleoli[262] and Cajal body[263], which are responsible for ribosomal RNA[264] and spliceosome maturation[265], respectively. Lastly, an

increase in SGs and a decrease in P-bodies is observed in neurons expressing RAN peptides[266]; in this case, RAN peptides may act similarly to small proteins such as NoBody[195] that dissolve P-bodies, releasing unstable RNAs to be sequestered by SGs. Additional studies are required to determine the effect of RAN peptides on RNA stability, P-body dynamics, and global RNA homeostasis.

2.4.3 RNA Transport

The diverse functions of RNA are determined, in part, by its subcellular localization. As a result, RNA transport mechanisms are crucial for RNA function, particularly in highly compartmentalized and morphologically complex cells such as neurons. Among the most important of these mechanisms is nucleocytoplasmic transport, in which RNA transcripts are shuttled from the nucleus to the cytoplasm. Several neurodegenerative diseases exhibit deficits in nucleocytoplasmic RNA transport, leading to RNA sequestration in the nucleus and widespread dysregulation of gene expression. Thus, interruption of nuclear export machinery can have severe consequences on neuronal health.

2.4.3.1 Impaired Nuclear Export

Nuclear mRNA export is triggered by deposition of the highly conserved translation export (TREX) complex at the 5' end of the nascent transcript[267]. The core of this complex, THO, recruits ALYREF and several other nuclear export factors[268–271]. ALYREF then binds to nuclear export factor 1 (NXF1)[272], triggering a shift from a conformation with low RNA binding affinity to one that readily binds the transcript[273,274]. NXF1 directs the transcript to the nuclear pore complex (NPC), a large multimeric structure that spans the nuclear envelope and enables the transport of

molecules into and out of the nucleus. NXF1 facilitates NPC docking and transcript translocation via interactions with NPC components containing low complexity domains enriched in phenylalanine and glycine residues[275].

Disruption of this pathway leads to nuclear retention of RNA, and which is then rapidly degraded by the nuclear exosome[276,277]. Interrupting nuclear RNA export can have severe consequences for neuronal survival, and mutations in nuclear export components are linked to several neurological and neurodevelopmental disorders. Chromosomal translocation and inactivation of THOC2, a subunit of the core TREX complex, leads to cognitive impairment, cerebellar hypoplasia, and congenital ataxia in humans[278]. Additionally, missense mutations in THOC2 have been implicated in fragile X syndrome[279], and mutations in a second THO subunit, THOC6, lead to intellectual disabilities[280]. Moreover, loss of function mutations in *Gle1* results in ALS[281] and fetal motor neuron disease[282]. Gle1 is a nuclear export mediator located on the cytoplasmic face of the nuclear pore that facilitates both the release of the transcript from the nuclear pore and its dissociation from export adaptor proteins[283], freeing it to undergo translation. This process may be specific to mRNAs with poly(A) tails, as depletion of Gle1 results in a nuclear accumulation and subsequent degradation of polyadenylated mRNAs[284,285].

Abnormal nucleocytoplasmic transport is also a characteristic finding in models of ALS[286–288], DM1[289] and HD[290,291]. Toxicity in these models can be suppressed by pharmacologic or genetic modulation of nuclear transport components, testifying to the broad significance of this pathway in disease pathogenesis. Moreover, age is a likely contributor to impaired nuclear import, as aged cells display abnormal NPCs and reduced expression of nucleocytoplasmic transport genes[292,293]; the resulting reduced fidelity in nuclear import/export is consistent with the observed age-dependent risk of nearly every neurodegenerative disease.

2.4.3.2 Disruption of the Nuclear Pore

In addition to disruption of the recruitment of the transcript to the pore, interruption of the pore itself can alter nucleocytoplasmic transport. RAN translation of repeat expansion mutations produces several DPRs. Some of these DPRs, including arginine-rich dipeptides generated from RAN translation of the *C9orf72* G_4C_2 repeat in familial ALS/FTD, clog the nuclear pore and inhibit the transport of RNA and other macromolecules into and out of the nucleus[294]. Again, this contributes to the nuclear retention of RNAs that are susceptible to exosome-mediated decay[276,277]. Arginine-containing DPRs are among the most toxic of the dipeptides in ALS/FTD models[262,295], suggesting that impaired nucleocytoplasmic transport contributes significantly to neurodegeneration in these disorders.

2.4.4 The RNA Exosome Complex

The exosome complex is an RNA degradation mechanism that contributes broadly to RNA turnover, surveillance, and processing. This complex works closely with other pathways to orchestrate the degradation of immature, abnormal, or misplaced RNA.

2.4.4.1 Exosome-Associated Mutations in Neurodegenerative Disease

Due to the importance of the exosome in regulating RNA decay, mutations in this complex can have severe implications. Mutations in *EXOSC3*, the gene encoding the core exosome component RRP40, are linked to autosomal recessive pontocerebellar hypoplasia type 1 (PCH1)[296]. This progressive neurodegenerative disease is characterized by atrophy of the pons and cerebellum and loss of spinal motor neurons, accompanied by developmental delay, muscle

atrophy, and difficulty breathing[297]. 37% of PCH1 patients exhibit *EXOSC3* mutations, most of which are heterozygous missense mutations[297]. Disease severity correlates with genotype, as patients with homozygous missense mutation fare better and those with a combined missense and null mutation fare worse[298].

Similarly, mutations in a gene encoding a separate exosome component, *EXOSC8*, result in cerebellar hypoplasia (CH)[299]. This autosomal recessive disorder is also characterized by progressive degeneration of the cerebellum, pons, and spinal motor neurons, as well as abnormal myelination. Though the mechanism is unclear, an increase in exosome substrates, including ARE-containing mRNAs encoding myelin proteins, in CH models suggests that impaired exosome function may contribute to dysmyelination of the involved tracts and subsequent neurodegeneration[299].

2.4.5 Alternative Splicing

Between 92 and 94% of all genes in the human genome are alternatively spliced[300], and the brain expresses more alternatively spliced genes than any other organ[301,302]. This suggests that alternative splicing is a key regulator of transcript stability and gene expression, and its misregulation can have severe effects on neuronal health[303].

2.4.5.1 Nonsense-Mediated Decay and Unannotated or "Cryptic" Exon Splicing

A primary consequence of alternative splicing is RNA destabilization[101]. As discussed above, in many cases alternative splicing may serve to regulate normal transcript levels. This is supported by the fact that over one third of RNA transcripts are spliced to include PTCs, and these transcripts are likely targeted for degradation via NMD[101]. Mutations that affect splicing and result

in either the inclusion of PTC-encoding exons or a shift the reading frame that uncovers 'silent' PTCs may destabilize transcripts and lead to disease via gene haploinsufficiency. For example, disease-associated missense *GRN* mutations cause ALS and FTD by altering mRNA splicing, triggering NMD of *GRN* transcripts, and consequent reductions in progranulin protein expression[304–307]. In other cases, mutations that create novel splice sites or the dysregulation of splicing factors leads to the inclusion of unannotated or "cryptic" exons and the production of a faulty transcripts that are eventually targeted for decay. Several regulatory proteins suppress these unannotated exon splicing events, including TDP43. Depletion of TDP43 results in a widespread increase in cryptic exon splicing events, and the inclusion of these exons may lead to NMD[308,309]. Many of these events are specific to neurons[310], which suggests that the disruption of TDP43-mediated cryptic exon regulation may contribute to ALS and FTD.

NMD can be manipulated through the modulation of specific pathway components: overexpression of UPF1 and UPF3B stimulate NMD, while UPF1 knockdown or the overexpression of UPF3A, an antagonistic paralog of UPF3B that sequesters UPF2, suppress NMD[311]. Consistent with a potential link between NMD and ALS/FTD pathogenesis, overexpression of UPF1 or UPF2 prevents FUS- and TDP43-mediated neurodegeneration in model systems[312]. One possibility is that UPF1 overexpression in these models prevents cell death by boosting endogenous NMD, thereby enabling the pathway to properly metabolize an overabundance of NMD substrates. However, further investigation is required to confirm and extend these findings.

2.4.6 Retrotransposons

Transposable elements (TEs) are mobile genetic elements that constitute a large portion of most eukaryotic genomes. Retrotransposons, which encode a reverse transcriptase and an integrase that allow them to "copy and paste" themselves from one region to another, represent approximately 40% of the human genome[313]. Though the vast majority of retrotransposons are inactive[314], some retain the ability to mobilize. Retrotransposition occurs approximately once every 10-100 births[315], and the insertion of these elements near or within active genes is a significant source of genomic instability and cellular toxicity[316,317]. Though transcription of these regions is downregulated[318,319], the transcripts that are transcribed are degraded via NMD[320] and other non-canonical pathways[321]. Several mechanisms have also evolved to suppress retrotransposon expression and prevent the resultant large scale deletions and genomic rearrangements[322], and the efficiency of these mechanisms declines with age[316,323,324]. Moreover, the elevated expression of retrotransposons correlates with several neurodegenerative disorders[325–327], suggesting that a reduction in retrotransposon repression may contribute to disease pathogenesis.

2.4.6.1 Retrotransposons in ALS

As previously discussed, TDP43 aggregation and mislocalization play a fundamental role in ALS and FTD, and TDP43 serves as a key regulator of alternative splicing for hundreds of transcripts. TDP43 also recognizes several TE-derived RNA transcripts[328], and this binding is reduced in FTD patients coincident with elevated TE expression. This suggests that TDP43 normally regulates TE expression, and the loss of functional TDP43 in FTD results in TE overexpression[328]. This is further supported by the finding that TEs are derepressed in ALS/FTD

models involving TDP43 overexpression or knockdown[328,329], suggesting that TE dysregulation may contribute to neurodegeneration in ALS and FTD. This may occur through activation of DNA damage-mediated programmed cell death due to the large scale deletions and genomic rearrangements that result from de-repressed TEs[329], and there is some evidence to suggest that TDP43 pathology impairs siRNA-mediated gene silencing, an essential system that normally protects the genome from retrotransposons[329]

Human endogenous retroviruses (HERVs) represent a subclass of retrotransposons originating from ancient viral infections that resulted in the integration of viral DNA into the host genome. The most recent of the retroviruses to integrate into the human genome is HERV-K[330]. The HERV-K envelope protein is expressed in both cortical and spinal neurons of ALS patients, suggesting activation of the retrovirus in disease. Furthermore, ectopic expression of the HERV-K envelope protein triggers neurodegeneration and motor dysfunction in mice[331]. Like other retrotransposons, HERV-K is regulated by TDP43, suggesting that HERV-K derepression in TDP43-deficient cells might contribute to neurodegeneration in ALS[331].

2.4.6.2 Retrotransposons in Aging

Age is a major risk factor for most neurodegenerative diseases, likely due to a reduced ability regulate protein degradation[332], oxidative stress[333], and DNA damage[334]. While retrotransposons are a significant source of genomic instability, additional evidence suggests that they are more destructive in aging brains. The expression and mobility of several TEs increase with advanced age[316,324]; these changes, in turn, are linked to progressive, age-dependent memory impairment and shortened lifespan[324]. Thus, the derepression of retrotransposons during normal aging could contribute to the age-related increase in risk for neurodegenerative diseases.

2.5 Conclusions and Future Directions

Neurodegenerative diseases vary widely in clinical presentation, neuropathology, and genetic background. However, it is becoming increasingly clear that alterations in RNA turnover are a key contributor to disease pathogenesis. The magnitude and extent of RNA dyshomeostasis observed in neurodegenerative disease models strongly suggests a fundamental disruption of one or more of the many mechanisms that tightly regulate RNA stability. While compensatory pathways may allow cells to cope with subtle changes in SG dynamics, alternative RNA splicing or RNA degradation, over time such pathways become less efficient and the ability of the cell to maintain RNA homeostasis slowly erodes. Mitotic cells evade toxicity by dilution and division, but for long-lived cells such as neurons, the resulting abnormalities eventually lead to cell death. Because altered RNA stability results from the disruption of several related but distinct pathways, it is unlikely that focusing on single transcripts will result in a cure. Instead, a more complete understanding of RNA degradation in both healthy and diseased conditions may highlight common mechanisms and key upstream elements that could be rationally targeted for therapeutic development.

2.6 Acknowledgements

Funding for this work was provided by the National Institutes of Health / National Institute for Neurological Disorders and Stroke (NIH/NINDS) R01NS097542 (SB), and Ann Arbor Active Against ALS (KW).

References

1. Yang, L., Duff, M. O., Graveley, B. R., Carmichael, G. G. & Chen, L.-L. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16 (2011).
2. Guhaniyogi, J. & Brewer, G. Regulation of mRNA stability in mammalian cells. *Gene* **265**, 11–23 (2001).
3. Gerstel, B., Tuite, M. F. & McCarthy, J. E. G. The effects of 5′-capping, 3′-polyadenylation and leader composition upon the translation and stability of mRNA in a cell-free extract derived from the yeast Saccharomyces cerevisiae. *Mol. Microbiol.* **6**, 2339–2348 (1992).
4. Huang, Y. & Carmichael, G. G. Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.* **16**, 1534–1542 (1996).
5. Bienroth, S., Keller, W. & Wahle, E. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* **12**, 585–594 (1993).
6. Wahle, E. Poly(A) Tail Length Control Is Caused by Termination of Processive Synthesis. *J. Biol. Chem.* **270**, 2800–2808 (1995).
7. Coller, J. M., Gray, N. K. & Wickens, M. P. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev.* **12**, 3226–3235 (1998).
8. Tian, B., Hu, J., Zhang, H. & Lutz, C. S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**, 201–212 (2005).
9. Shell, S. A., Hesse, C., Morris, S. M., Jr & Milcarek, C. Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *J. Biol. Chem.* **280**, 39950–39961 (2005).
10. Wood, A. J. *et al.* Regulation of alternative polyadenylation by genomic imprinting. *Genes Dev.* **22**, 1141–1146 (2008).
11. Danckwardt, S. *et al.* Splicing factors stimulate polyadenylation via USEs at non-canonical 3′ end formation signals. *EMBO J.* **26**, 2658–2669 (2007).
12. Hall-Pogar, T., Liang, S., Hague, L. K. & Lutz, C. S. Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. *RNA* **13**, 1103–1115 (2007).
13. Tian, B., Pan, Z. & Lee, J. Y. Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. *Genome Res.* **17**, 156–165 (2007).
14. Meyer, S., Temme, C. & Wahle, E. Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39**, 197–216 (2004).
15. Weidmann, C. A., Raynard, N. A., Blewett, N. H., Van Etten, J. & Goldstrohm, A. C. The RNA binding domain of Pumilio antagonizes poly-adenosine binding protein and accelerates deadenylation. *RNA* **20**, 1298–1319 (2014).
16. Furuichi, Y., LaFiandra, A. & Shatkin, A. J. 5'-Terminal structure and mRNA stability. *Nature* **266**, 235–239 (1977).
17. Shimotohno, K., Kodama, Y., Hashimoto, J. & Miura, K. I. Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 2734–2738 (1977).
18. Murthy, K. G. K., Park, P. & Manley, J. L. A nuclear micrococcal-sensitive, ATP-dependent exoribonuclease degrades uncapped but not capped RNA substratesx. *Nucleic Acids Res.* **19**, 2685–2692 (1991).

19. Muthukrishnan, S., Both, G. W., Furuichi, Y. & Shatkin, A. J. 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* **255**, 33–37 (1975).
20. Gillian-Daniel, D. L., Gray, N. K., Aström, J., Barkoff, A. & Wickens, M. Modifications of the 5' cap of mRNAs during Xenopus oocyte maturation: independence from changes in poly(A) length and impact on translation. *Mol. Cell. Biol.* **18**, 6152–6163 (1998).
21. Konarska, M. M., Padgett, R. A. & Sharp, P. A. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38**, 731–736 (1984).
22. Edery, I. & Sonenberg, N. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7590–7594 (1985).
23. Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W. & Gilmartin, G. M. Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11893–11898 (1997).
24. Jarmolowski, A., Boelens, W. C., Izaurralde, E. & Mattaj, I. W. Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell Biol.* **124**, 627–635 (1994).
25. Fresco, L. D. & Buratowski, S. Conditional mutants of the yeast mRNA capping enzyme show that the cap enhances, but is not required for, mRNA splicing. *RNA* **2**, 584–596 (1996).
26. Glover-Cutter, K., Kim, S., Espinosa, J. & Bentley, D. L. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.* **15**, 71–78 (2008).
27. Shatkin, A. J. Capping of eucaryotic mRNAs. *Cell* **9**, 645–653 (1976).
28. Nojima, T., Hirose, T., Kimura, H. & Hagiwara, M. The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *J. Biol. Chem.* **282**, 15645–15651 (2007).
29. Cheng, H. *et al.* Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127**, 1389–1400 (2006).
30. Sato, H. & Maquat, L. E. Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. *Genes Dev.* **23**, 2537–2550 (2009).
31. Dias, S. M. G., Wilson, K. F., Rojas, K. S., Ambrosio, A. L. B. & Cerione, R. A. The molecular basis for the regulation of the cap-binding complex by the importins. *Nat. Struct. Mol. Biol.* **16**, 930–937 (2009).
32. Schwartz, D. C. & Parker, R. mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Mol. Cell. Biol.* **20**, 7933–7942 (2000).
33. Grudzien, E., Kalek, M., Jemielity, J., Darzynkiewicz, E. & Rhoads, R. E. Differential inhibition of mRNA degradation pathways by novel cap analogs. *J. Biol. Chem.* **281**, 1857–1867 (2006).
34. Jiao, X., Chang, J. H., Kilic, T., Tong, L. & Kiledjian, M. A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol. Cell* **50**, 104–115 (2013).
35. Braun, J. E. *et al.* A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. *Nat. Struct. Mol. Biol.* **19**, 1324–1331 (2012).
36. Varani, G. Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 379–404 (1995).
37. Emory, S. A., Bouvet, P. & Belasco, J. G. A 5'-terminal stem-loop structure can stabilize

mRNA in Escherichia coli. *Genes Dev.* **6**, 135–148 (1992).
38. Hambraeus, G., Karhumaa, K. & Rutberg, B. A 5′ stem–loop and ribosome binding but not translation are important for the stability of Bacillus subtilis aprE leader mRNA. *Microbiology* **148**, 1795–1803 (2002).
39. Higgs, D. C., Shapiro, R. S., Kindle, K. L. & Stern, D. B. Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. *Mol. Cell. Biol.* **19**, 8479–8491 (1999).
40. Zou, Z., Eibl, C. & Koop, H.-U. The stem-loop region of the tobacco psbA 5′ UTR is an important determinant of mRNA stability and translation efficiency. *Mol. Genet. Genomics* **269**, 340–349 (2003).
41. Muslimov, I. A. *et al.* Dendritic Transport and Localization of Protein Kinase Mζ mRNA: IMPLICATIONS FOR MOLECULAR MEMORY CONSOLIDATION. *J. Biol. Chem.* **279**, 52613–52622 (2004).
42. Chao, J. A. *et al.* ZBP1 recognition of β-actin zipcode induces RNA looping. *Genes Dev.* **24**, 148–158 (2010).
43. Kim, H. H., Lee, S. J., Gardiner, A. S., Perrone-Bizzozero, N. I. & Yoo, S. Different motif requirements for the localization zipcode element of β-actin mRNA binding by HuD and ZBP1. *Nucleic Acids Res.* **43**, 7432–7446 (2015).
44. Kadrmas, J. L., Ravin, A. J. & Leontis, N. B. Relative stabilities of DNA three-way, four-way and five-way junctions (multi-helix junction loops): unpaired nucleotides can be stabilizing or destabilizing. *Nucleic Acids Res.* **23**, 2212–2222 (1995).
45. Ke, A., Zhou, K., Ding, F., Cate, J. H. D. & Doudna, J. A. A conformational switch controls hepatitis delta virus ribozyme catalysis. *Nature* **429**, 201–205 (2004).
46. Rastogi, T., Beattie, T. L., Olive, J. E. & Collins, R. A. A long-range pseudoknot is required for activity of the Neurospora VS ribozyme. *EMBO J.* **15**, 2820–2825 (1996).
47. Theimer, C. A., Blois, C. A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell* **17**, 671–682 (2005).
48. Tang, C. K. & Draper, D. E. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell* **57**, 531–536 (1989).
49. Koritzinsky, M. *et al.* Phosphorylation of eIF2alpha is required for mRNA translation inhibition and survival during moderate hypoxia. *Radiother. Oncol.* **83**, 353–361 (2007).
50. Spriggs, K. A., Bushell, M. & Willis, A. E. Translational regulation of gene expression during conditions of cell stress. *Mol. Cell* **40**, 228–237 (2010).
51. Kedersha, N. L., Gupta, M., Li, W., Miller, I. & Anderson, P. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J. Cell Biol.* **147**, 1431–1442 (1999).
52. Harding, H. P. *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* **6**, 1099–1108 (2000).
53. Mazroui, R. *et al.* Inhibition of ribosome recruitment induces stress granule formation independently of eukaryotic initiation factor 2alpha phosphorylation. *Mol. Biol. Cell* **17**, 4212–4219 (2006).
54. Kedersha, N. *et al.* Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol. Biol. Cell* **13**, 195–210 (2002).
55. Kimball, S. R., Horetsky, R. L., Ron, D., Jefferson, L. S. & Harding, H. P. Mammalian

stress granules represent sites of accumulation of stalled translation initiation complexes. *Am. J. Physiol. Cell Physiol.* **284**, C273–84 (2003).
56. Buchan, J. R., Yoon, J.-H. & Parker, R. Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. *J. Cell Sci.* **124**, 228–239 (2011).
57. Kwon, S., Zhang, Y. & Matthias, P. The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev.* **21**, 3381–3394 (2007).
58. Gallouzi, I. E. *et al.* A novel phosphorylation-dependent RNase activity of GAP-SH3 binding protein: a potential link between signal transduction and RNA stability. *Mol. Cell. Biol.* **18**, 3956–3965 (1998).
59. Schmidlin, M. *et al.* The ARE-dependent mRNA-destabilizing activity of BRF1 is regulated by protein kinase B. *EMBO J.* **23**, 4760–4769 (2004).
60. Tourrière, H. *et al.* RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization. *Mol. Cell. Biol.* **21**, 7747–7760 (2001).
61. Ohn, T., Kedersha, N., Hickman, T., Tisdale, S. & Anderson, P. A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **10**, 1224–1231 (2008).
62. Gilks, N. *et al.* Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**, 5383–5398 (2004).
63. Colombrita, C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051–1061 (2009).
64. Buchan, J. R. & Parker, R. Eukaryotic stress granules: the ins and outs of translation. *Mol. Cell* **36**, 932–941 (2009).
65. Laroia, G., Cuesta, R., Brewer, G. & Schneider, R. J. Control of mRNA decay by heat shock-ubiquitin-proteasome pathway. *Science* **284**, 499–502 (1999).
66. Gowrishankar, G. *et al.* Inhibition of mRNA deadenylation and degradation by different types of cell stress. *Biol. Chem.* **387**, 323–327 (2006).
67. Hilgers, V., Teixeira, D. & Parker, R. Translation-independent inhibition of mRNA deadenylation during stress in Saccharomyces cerevisiae. *RNA* **12**, 1835–1845 (2006).
68. Thomas, M. G., Martinez Tosar, L. J., Desbats, M. A., Leishman, C. C. & Boccaccio, G. L. Mammalian Staufen 1 is recruited to stress granules and impairs their assembly. *J. Cell Sci.* **122**, 563–573 (2009).
69. Tsai, N.-P., Ho, P.-C. & Wei, L.-N. Regulation of stress granule dynamics by Grb7 and FAK signalling pathway. *EMBO J.* **27**, 715–726 (2008).
70. Rikhvanov, E. G., Romanova, N. V. & Chernoff, Y. O. Chaperone effects on prion and nonprion aggregates. *Prion* **1**, 217–222 (2007).
71. Raijmakers, R., Egberts, W. V., van Venrooij, W. J. & Pruijn, G. J. M. Protein–Protein Interactions between Human Exosome Components Support the Assembly of RNase PH-type Subunits into a Six-membered PNPase-like Ring. *J. Mol. Biol.* **323**, 653–663 (2002).
72. Schilders, G., van Dijk, E., Raijmakers, R. & Pruijn, G. J. M. Cell and Molecular Biology of the Exosome: How to Make or Break an RNA. in *International Review of Cytology* **251**, 159–208 (Academic Press, 2006).
73. Budenholzer, L., Cheng, C. L., Li, Y. & Hochstrasser, M. Proteasome Structure and Assembly. *J. Mol. Biol.* (2017). doi:10.1016/j.jmb.2017.05.027
74. Houseley, J., LaCava, J. & Tollervey, D. RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.* **7**, 529–539 (2006).
75. Chen, C. Y. *et al.* AU binding proteins recruit the exosome to degrade ARE-containing

mRNAs. *Cell* **107**, 451–464 (2001).
76. Kowalinski, E., Schuller, A., Green, R. & Conti, E. Saccharomyces cerevisiae Ski7 Is a GTP-Binding Protein Adopting the Characteristic Conformation of Active Translational GTPases. *Structure* **23**, 1336–1343 (2015).
77. Frischmeyer, P. A. *et al.* An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* **295**, 2258–2261 (2002).
78. Allmang, C. *et al.* Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399–5410 (1999).
79. Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199 (1998).
80. Thermann, R. *et al.* Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* **17**, 3484–3494 (1998).
81. Kashima, I. *et al.* Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **20**, 355–367 (2006).
82. Melero, R. *et al.* Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. *Structure* **22**, 1105–1119 (2014).
83. Deniaud, A. *et al.* A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic Acids Res.* **43**, 7600–7611 (2015).
84. Franks, T. M., Singh, G. & Lykke-Andersen, J. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. *Cell* **143**, 938–950 (2010).
85. Fiorini, F., Bagchi, D., Le Hir, H. & Croquette, V. Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nat. Commun.* **6**, 7581 (2015).
86. Huntzinger, E., Kashima, I., Fauser, M., Saulière, J. & Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14**, 2609–2617 (2008).
87. Eberle, A. B., Lykke-Andersen, S., Mühlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* **16**, 49–55 (2009).
88. Loh, B., Jonas, S. & Izaurralde, E. The SMG5–SMG7 heterodimer directly recruits the CCR4–NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* **27**, 2125–2138 (2013).
89. Unterholzner, L. & Izaurralde, E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* **16**, 587–596 (2004).
90. Hogg, J. R. & Goff, S. P. Upf1 senses 3′ UTR length to potentiate mRNA decay. *Cell* **143**, 379–389 (2010).
91. Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A. & Mühlemann, O. Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim. Biophys. Acta* **1829**, 612–623 (2013).
92. Avery, P. *et al.* Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. *RNA* **17**, 624–638 (2011).
93. Lou, C. H. *et al.* Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. *Cell Rep.* **6**, 748–764 (2014).
94. Gloggnitzer, J. *et al.* Nonsense-mediated mRNA decay modulates immune receptor levels

to regulate plant antibacterial defense. *Cell Host Microbe* **16**, 376–390 (2014).
95. Gardner, L. B. Nonsense-mediated RNA decay regulation by cellular stress: implications for tumorigenesis. *Mol. Cancer Res.* **8**, 295–308 (2010).
96. Balistreri, G. *et al.* The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. *Cell Host Microbe* **16**, 403–411 (2014).
97. Giorgi, C. *et al.* The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. *Cell* **130**, 179–191 (2007).
98. Colak, D., Ji, S.-J., Porse, B. T. & Jaffrey, S. R. Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell* **153**, 1252–1265 (2013).
99. Weg-Remers, S., Ponta, H., Herrlich, P. & König, H. Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.* **20**, 4194–4203 (2001).
100. van der Houven van Oordt, W. *et al.* The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J. Cell Biol.* **149**, 307–316 (2000).
101. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 189–192 (2003).
102. Eom, T. *et al.* NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife* **2**, e00178 (2013).
103. Winter, J. *et al.* Regulation of the MID1 protein function is fine-tuned by a complex pattern of alternative splicing. *Hum. Genet.* **114**, 541–552 (2004).
104. Sureau, A., Gattoni, R., Dooghe, Y., Stévenin, J. & Soret, J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J.* **20**, 1785–1796 (2001).
105. Wilson, G. M. *et al.* Regulation of AUF1 expression via conserved alternatively spliced elements in the 3' untranslated region. *Mol. Cell. Biol.* **19**, 4056–4064 (1999).
106. Lamba, J. K. *et al.* Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons. *Hum. Mol. Genet.* **12**, 99–109 (2003).
107. Jones, R. B. *et al.* The nonsense-mediated decay pathway and mutually exclusive expression of alternatively spliced FGFR2IIIb and -IIIc mRNAs. *J. Biol. Chem.* **276**, 4158–4167 (2001).
108. Lareau, L. F., Brooks, A. N., Soergel, D. A. W., Meng, Q. & Brenner, S. E. The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv. Exp. Med. Biol.* **623**, 190–211 (2007).
109. Morris, D. R. & Geballe, A. P. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **20**, 8635–8642 (2000).
110. Kozak, M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867–19870 (1991).
111. Matsui, M., Yachie, N., Okada, Y., Saito, R. & Tomita, M. Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse. *FEBS Lett.* **581**, 4184–4188 (2007).
112. Calvo, S. E., Pagliarini, D. J. & Mootha, V. K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7507–7512 (2009).
113. Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes

genomic noise. *Nat. Genet.* **36**, 1073–1078 (2004).
114. Ramani, A. K. *et al.* High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. *Genome Biol.* **10**, R101 (2009).
115. He, F. *et al.* Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5’ to 3' mRNA decay pathways in yeast. *Mol. Cell* **12**, 1439–1452 (2003).
116. van Hoof, A., Frischmeyer, P. A., Dietz, H. C. & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262–2264 (2002).
117. Karzai, A. W., Roche, E. D. & Sauer, R. T. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat. Struct. Biol.* **7**, 449–455 (2000).
118. Benard, L., Carroll, K., Valle, R. C., Masison, D. C. & Wickner, R. B. The ski7 antiviral protein is an EF1-alpha homolog that blocks expression of non-Poly(A) mRNA in Saccharomyces cerevisiae. *J. Virol.* **73**, 2893–2900 (1999).
119. Anderson, J. S. & Parker, R. P. The 3’ to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.* **17**, 1497–1506 (1998).
120. Caponigro, G. & Parker, R. Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev.* **9**, 2421–2432 (1995).
121. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006).
122. Tollervey, D. Molecular biology: RNA lost in translation. *Nature* **440**, 425–426 (2006).
123. Clement, S. L. & Lykke-Andersen, J. No mercy for messages that mess with the ribosome. *Nat. Struct. Mol. Biol.* **13**, 299–301 (2006).
124. Young, S. K., Palam, L. R., Wu, C., Sachs, M. S. & Wek, R. C. Ribosome Elongation Stall Directs Gene-specific Translation in the Integrated Stress Response. *J. Biol. Chem.* **291**, 6546–6558 (2016).
125. Passos, D. O. *et al.* Analysis of Dom34 and its function in no-go decay. *Mol. Biol. Cell* **20**, 3025–3032 (2009).
126. Graille, M., Chaillet, M. & van Tilbeurgh, H. Structure of yeast Dom34: a protein related to translation termination factor Erf1 and involved in No-Go decay. *J. Biol. Chem.* **283**, 7145–7154 (2008).
127. 이형호. Structural and functional insights into Dom34, a key component of No-Go mRNA decay, and structure of a metal Ion-Bound IS200 transposase. (서울대학교 대학원, 2007).
128. Carr-Schmid, A., Pfund, C., Craig, E. A. & Kinzy, T. G. Novel G-protein complex whose requirement is linked to the translational status of the cell. *Mol. Cell. Biol.* **22**, 2564–2574 (2002).
129. Protzel, A. & Morris, A. J. Gel chromatographic analysis of nascent globin chains. Evidence of nonuniform size distribution. *J. Biol. Chem.* **249**, 4594–4600 (1974).
130. Beelman, C. A. & Parker, R. Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *J. Biol. Chem.* **269**, 9687–9692 (1994).
131. Nagai, K. *et al.* Structure, function and evolution of the signal recognition particle. *EMBO J.* **22**, 3479–3485 (2003).
132. Wang, Z. & Sachs, M. S. Ribosome stalling is responsible for arginine-specific translational attenuation in Neurospora crassa. *Mol. Cell. Biol.* **17**, 4904–4913 (1997).
133. Chen, C. Y. & Shyu, A. B. AU-rich elements: characterization and importance in mRNA

degradation. *Trends Biochem. Sci.* **20**, 465–470 (1995).
134. Shaw, G. & Kamen, R. A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46**, 659–667 (1986).
135. Loflin, P., Chen, C. Y. & Shyu, A. B. Unraveling a cytoplasmic role for hnRNP D in the in vivo mRNA destabilization directed by the AU-rich element. *Genes Dev.* **13**, 1884–1897 (1999).
136. Sarkar, B., Xi, Q., He, C. & Schneider, R. J. Selective degradation of AU-rich mRNAs promoted by the p37 AUF1 protein isoform. *Mol. Cell. Biol.* **23**, 6685–6693 (2003).
137. Lal, A. *et al.* Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO J.* **23**, 3092–3102 (2004).
138. Raineri, I., Wegmueller, D., Gross, B., Certa, U. & Moroni, C. Roles of AUF1 isoforms, HuR and BRF1 in ARE-dependent mRNA turnover studied by RNA interference. *Nucleic Acids Res.* **32**, 1279–1288 (2004).
139. Carballo, E., Gilkeson, G. S. & Blackshear, P. J. Bone marrow transplantation reproduces the tristetraprolin-deficiency syndrome in recombination activating gene-2 (-/-) mice. Evidence that monocyte/macrophage progenitors may be responsible for TNFalpha overproduction. *J. Clin. Invest.* **100**, 986–995 (1997).
140. Carballo, E., Lai, W. S. & Blackshear, P. J. Evidence that tristetraprolin is a physiological regulator of granulocyte-macrophage colony-stimulating factor messenger RNA deadenylation and stability. *Blood* **95**, 1891–1899 (2000).
141. Ogilvie, R. L. *et al.* Tristetraprolin down-regulates IL-2 gene expression through AU-rich element-mediated mRNA decay. *J. Immunol.* **174**, 953–961 (2005).
142. Carballo, E., Lai, W. S. & Blackshear, P. J. Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. *Science* **281**, 1001–1005 (1998).
143. Lai, W. S. & Blackshear, P. J. Interactions of CCCH zinc finger proteins with mRNA: tristetraprolin-mediated AU-rich element-dependent mRNA degradation can occur in the absence of a poly(A) tail. *J. Biol. Chem.* **276**, 23144–23154 (2001).
144. Lai, W. S., Kennington, E. A. & Blackshear, P. J. Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease. *Mol. Cell. Biol.* **23**, 3798–3812 (2003).
145. Lai, W. S. *et al.* Evidence that tristetraprolin binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA. *Mol. Cell. Biol.* **19**, 4311–4323 (1999).
146. Sawaoka, H., Dixon, D. A., Oates, J. A. & Boutaud, O. Tristetraprolin Binds to the 3′-Untranslated Region of Cyclooxygenase-2 mRNA: A POLYADENYLATION VARIANT IN A CANCER CELL LINE LACKS THE BINDING SITE. *J. Biol. Chem.* **278**, 13928–13935 (2003).
147. Stoecklin, G., Ming, X. F., Looser, R. & Moroni, C. Somatic mRNA turnover mutants implicate tristetraprolin in the interleukin-3 mRNA degradation pathway. *Mol. Cell. Biol.* **20**, 3753–3763 (2000).
148. Wilson, G. M. & Brewer, G. The search for trans-acting factors controlling messenger RNA decay. *Prog. Nucleic Acid Res. Mol. Biol.* **62**, 257–291 (1999).
149. Wilson, T. & Treisman, R. Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. *Nature* **336**, 396–399 (1988).
150. Shyu, A. B., Belasco, J. G. & Greenberg, M. E. Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes Dev.* **5**, 221–

231 (1991).
151. Brewer, G. & Ross, J. Poly(A) shortening and degradation of the 3' A+U-rich sequences of human c-myc mRNA in a cell-free system. *Mol. Cell. Biol.* **8**, 1697–1708 (1988).
152. Mukherjee, D. *et al.* The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *EMBO J.* **21**, 165–174 (2002).
153. Gherzi, R. *et al.* A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol. Cell* **14**, 571–583 (2004).
154. Kedersha, N. *et al.* Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**, 871–884 (2005).
155. Kedersha, N. & Anderson, P. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem. Soc. Trans.* **30**, 963–969 (2002).
156. Levy, N. S., Chung, S., Furneaux, H. & Levy, A. P. Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR. *J. Biol. Chem.* **273**, 6417–6423 (1998).
157. Peng, S. S., Chen, C. Y., Xu, N. & Shyu, A. B. RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.* **17**, 3461–3470 (1998).
158. Rodriguez-Pascual, F. *et al.* Complex contribution of the 3'-untranslated region to the expressional regulation of the human inducible nitric-oxide synthase gene. Involvement of the RNA-binding protein HuR. *J. Biol. Chem.* **275**, 26040–26049 (2000).
159. Dean, J. L. E. *et al.* The 3′ Untranslated Region of Tumor Necrosis Factor Alpha mRNA Is a Target of the mRNA-Stabilizing Factor HuR. *Mol. Cell. Biol.* **25**, 3400 (2005).
160. Kim, Y. K., Furic, L., Desgroseillers, L. & Maquat, L. E. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**, 195–208 (2005).
161. Kim, Y. K. *et al.* Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J.* **26**, 2670–2681 (2007).
162. Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284–288 (2011).
163. Gong, C., Kim, Y. K., Woeller, C. F., Tang, Y. & Maquat, L. E. SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs. *Genes Dev.* **23**, 54–66 (2009).
164. Cai, X., Hagedorn, C. H. & Cullen, B. R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957–1966 (2004).
165. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
166. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
167. Gregory, R. I., Chendrimada, T. P. & Shiekhattar, R. MicroRNA biogenesis: isolation and characterization of the microprocessor complex. *Methods Mol. Biol.* **342**, 33–47 (2006).
168. Murchison, E. P. & Hannon, G. J. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr. Opin. Cell Biol.* **16**, 223–229 (2004).
169. Lund, E. & Dahlberg, J. E. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 59–66 (2006).
170. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–216 (2003).
171. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208 (2003).

172. Hausser, J., Syed, A. P., Bilen, B. & Zavolan, M. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.* **23**, 604–615 (2013).
173. Fang, Z. & Rajewsky, N. The impact of miRNA target sites in coding sequences and in 3'UTRs. *PLoS One* **6**, e18067 (2011).
174. Alemán, L. M., Doench, J. & Sharp, P. A. Comparison of siRNA-induced off-target RNA and protein effects. *RNA* **13**, 385–395 (2007).
175. Eulalio, A., Huntzinger, E. & Izaurralde, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat. Struct. Mol. Biol.* **15**, 346–353 (2008).
176. Behm-Ansmant, I. *et al.* mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20**, 1885–1898 (2006).
177. Wu, L., Fan, J. & Belasco, J. G. MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4034–4039 (2006).
178. Stoecklin, G., Lu, M., Rattenbacher, B. & Moroni, C. A constitutive decay element promotes tumor necrosis factor alpha mRNA degradation via an AU-rich element-independent pathway. *Mol. Cell. Biol.* **23**, 3506–3515 (2003).
179. Leppek, K. *et al.* Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell* **153**, 869–881 (2013).
180. Tan, D., Zhou, M., Kiledjian, M. & Tong, L. The ROQ domain of Roquin recognizes mRNA constitutive-decay element and double-stranded RNA. *Nat. Struct. Mol. Biol.* **21**, 679–685 (2014).
181. Marzluff, W. F., Wagner, E. J. & Duronio, R. J. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* **9**, 843–854 (2008).
182. Mullen, T. E. & Marzluff, W. F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev.* **22**, 50–65 (2008).
183. Kaygun, H. & Marzluff, W. F. Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat. Struct. Mol. Biol.* **12**, 794–800 (2005).
184. Eulalio, A., Behm-Ansmant, I., Schweizer, D. & Izaurralde, E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol. Cell. Biol.* **27**, 3970–3981 (2007).
185. Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M. & Parker, R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371–382 (2005).
186. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat. Cell Biol.* **7**, 719–723 (2005).
187. Brengues, M., Teixeira, D. & Parker, R. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* **310**, 486–489 (2005).
188. Koritzinsky, M. *et al.* Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control. *EMBO J.* **25**, 1114–1125 (2006).
189. Coller, J. & Parker, R. General translational repression by activators of mRNA decapping. *Cell* **122**, 875–886 (2005).
190. Parker, R. & Sheth, U. P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635–646 (2007).
191. Holmes, L. E. A., Campbell, S. G., De Long, S. K., Sachs, A. B. & Ashe, M. P. Loss of

translational control in yeast compromised for the major mRNA decay pathway. *Mol. Cell. Biol.* **24**, 2998–3010 (2004).
192. Sheth, U. & Parker, R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300**, 805–808 (2003).
193. Cougot, N., Babajko, S. & Séraphin, B. Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.* **165**, 31–40 (2004).
194. Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. & Filipowicz, W. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111–1124 (2006).
195. D'Lima, N. G. *et al.* A human microprotein that interacts with the mRNA decapping complex. *Nat. Chem. Biol.* **13**, 174–180 (2017).
196. Cougot, N. *et al.* Overexpression of MLN51 triggers P-body disassembly and formation of a new type of RNA granules. *J. Cell Sci.* **127**, 4692–4701 (2014).
197. Aulas, A. & Vande Velde, C. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? *Front. Cell. Neurosci.* **9**, 423 (2015).
198. Zinszner, H., Sok, J., Immanuel, D., Yin, Y. & Ron, D. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J. Cell Sci.* **110 (Pt 15)**, 1741–1750 (1997).
199. Åman, P. *et al.* Expression Patterns of the Human Sarcoma-Associated GenesFUSandEWSand the Genomic Structure ofFUS. *Genomics* **37**, 1–8 (1996).
200. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
201. Kwiatkowski, T. J., Jr *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).
202. Dormann, D. *et al.* ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *EMBO J.* **29**, 2841–2857 (2010).
203. Van Deerlin, V. M. *et al.* TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol.* **7**, 409–416 (2008).
204. Ayala, Y. M. *et al.* Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778–3785 (2008).
205. Zhang, Z. C. & Chook, Y. M. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12017–12021 (2012).
206. McDonald, K. K. *et al.* TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum. Mol. Genet.* **20**, 1400–1410 (2011).
207. Liu-Yesucevitz, L. *et al.* Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* **5**, e13250 (2010).
208. Parker, S. J. *et al.* Endogenous TDP-43 localized to stress granules can subsequently form protein aggregates. *Neurochem. Int.* **60**, 415–424 (2012).
209. 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).
210. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488–1497 (2012).
211. Voigt, A. *et al.* TDP-43-mediated neuron loss in vivo requires RNA-binding activity. *PLoS*

One **5**, e12247 (2010).
212. Daigle, J. G. *et al.* RNA-binding ability of FUS regulates neurodegeneration, cytoplasmic mislocalization and incorporation into stress granules associated with FUS carrying ALS-linked mutations. *Hum. Mol. Genet.* **22**, 1193–1205 (2013).
213. Johnson, J. O. *et al.* Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat. Neurosci.* **17**, 664–666 (2014).
214. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).
215. Mackenzie, I. R. *et al.* TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* **95**, 808–816.e9 (2017).
216. Bug, M. & Meyer, H. Expanding into new markets--VCP/p97 in endocytosis and autophagy. *J. Struct. Biol.* **179**, 78–82 (2012).
217. Dantuma, N. P., Acs, K. & Luijsterburg, M. S. Should I stay or should I go: VCP/p97-mediated chromatin extraction in the DNA damage response. *Exp. Cell Res.* **329**, 9–17 (2014).
218. Buchan, J. R., Kolaitis, R.-M., Taylor, J. P. & Parker, R. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* **153**, 1461–1474 (2013).
219. Koppers, M. *et al.* VCP mutations in familial and sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* **33**, 837.e7–13 (2012).
220. Treangen, T. J. & Salzberg, S. L. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* **13**, 36–46 (2011).
221. Echeverria, G. V. & Cooper, T. A. RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity. *Brain Res.* **1462**, 100–111 (2012).
222. Mohan, A., Goodwin, M. & Swanson, M. S. RNA-protein interactions in unstable microsatellite diseases. *Brain Res.* **1584**, 3–14 (2014).
223. Kiliszek, A. & Rypniewski, W. Structural studies of CNG repeats. *Nucleic Acids Res.* **42**, 8189–8199 (2014).
224. Napierała, M. & Krzyzosiak, W. J. CUG repeats present in myotonin kinase RNA form metastable 'slippery' hairpins. *J. Biol. Chem.* **272**, 31079–31085 (1997).
225. Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195–200 (2014).
226. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **34**, 5402–5415 (2006).
227. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. & Lipps, H. J. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat. Struct. Mol. Biol.* **12**, 847–854 (2005).
228. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
229. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
230. Rutherford, N. J. *et al.* Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. *Neurobiol. Aging* **33**, 2950.e5–7 (2012).
231. van Blitterswijk, M. *et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional

cohort study. *Lancet Neurol.* **12**, 978–988 (2013).
232. Fratta, P. *et al.* C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2**, 1016 (2012).
233. Reddy, K., Zamiri, B., Stanley, S. Y. R., Macgregor, R. B., Jr & Pearson, C. E. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* **288**, 9860–9866 (2013).
234. Reddy, K. *et al.* Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Res.* **39**, 1749–1762 (2011).
235. Lin, Y., Dent, S. Y. R., Wilson, J. H., Wells, R. D. & Napierala, M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 692–697 (2010).
236. Belotserkovskii, B. P., Mirkin, S. M. & Hanawalt, P. C. DNA sequences that interfere with transcription: implications for genome function and stability. *Chem. Rev.* **113**, 8620–8637 (2013).
237. Groh, M., Lufino, M. M. P., Wade-Martins, R. & Gromak, N. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.* **10**, e1004318 (2014).
238. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711–721 (2003).
239. Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **52**, 583–590 (2013).
240. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
241. Walker, C. *et al.* C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nat. Neurosci.* **20**, 1225–1235 (2017).
242. de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K. & Krzyzosiak, W. J. Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic Acids Res.* **39**, 3852–3863 (2011).
243. Conlon, E. G. *et al.* The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS patient brains. *Elife* **5**, e17820 (2016).
244. Warf, M. B. & Berglund, J. A. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA* **13**, 2238–2251 (2007).
245. Kino, Y. *et al.* Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.* **13**, 495–507 (2004).
246. Du, H. *et al.* Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.* **17**, 187–193 (2010).
247. Miller, J. W. *et al.* Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* **19**, 4439–4448 (2000).
248. Konieczny, P. *et al.* Myotonic dystrophy: candidate small molecule therapeutics. *Drug Discov. Today* (2017). doi:10.1016/j.drudis.2017.07.011
249. Taneja, K. L., McCurrach, M., Schalling, M., Housman, D. & Singer, R. H. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell*

Biol. **128**, 995–1002 (1995).
250. Liquori, C. L. *et al.* Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867 (2001).
251. Lu, X., Timchenko, N. A. & Timchenko, L. T. Cardiac elav-type RNA-binding protein (ETR-3) binds to RNA CUG repeats expanded in myotonic dystrophy. *Hum. Mol. Genet.* **8**, 53–60 (1999).
252. Mizielinska, S. *et al.* C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* **126**, 845–857 (2013).
253. Vatovec, S., Kovanda, A. & Rogelj, B. Unconventional features of C9ORF72 expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **35**, 2421.e1–2421.e12 (2014).
254. Zu, T. *et al.* Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 260–265 (2011).
255. Kearse, M. G. *et al.* CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol. Cell* **62**, 314–322 (2016).
256. Todd, P. K. *et al.* CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **78**, 440–455 (2013).
257. Bañez-Coronel, M. *et al.* RAN Translation in Huntington Disease. *Neuron* **88**, 667–677 (2015).
258. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E4968–77 (2013).
259. Zu, T. *et al.* RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron* **95**, 1292–1305.e5 (2017).
260. May, S. *et al.* C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol.* **128**, 485–503 (2014).
261. Tran, H. *et al.* Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron* **87**, 1207–1214 (2015).
262. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).
263. Lee, K.-H. *et al.* C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **167**, 774–788.e17 (2016).
264. Henras, A. K., Plisson-Chastang, C., O'Donohue, M.-F., Chakraborty, A. & Gleizes, P.-E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **6**, 225–242 (2015).
265. Staněk, D. *et al.* Spliceosomal Small Nuclear Ribonucleoprotein Particles Repeatedly Cycle through Cajal Bodies. *Mol. Biol. Cell* **19**, 2534–2543 (2008).
266. Wen, X. *et al.* Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* **84**, 1213–1225 (2014).
267. Masuda, S. *et al.* Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19**, 1512–1517 (2005).
268. Luo, M. L. *et al.* Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413**, 644–647 (2001).
269. Hautbergue, G. M. *et al.* UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.* **19**, 1918–1924 (2009).

270. Viphakone, N. *et al.* Luzp4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Res.* **43**, 2353–2366 (2015).
271. Chang, C.-T. *et al.* Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.* **32**, 473–486 (2013).
272. Hautbergue, G. M., Hung, M.-L., Golovanov, A. P., Lian, L.-Y. & Wilson, S. A. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5154–5159 (2008).
273. Viphakone, N. *et al.* TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.* **3**, 1006 (2012).
274. Stutz, F. *et al.* REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* **6**, 638–650 (2000).
275. Wickramasinghe, V. O. *et al.* mRNA export from mammalian cell nuclei is dependent on GANP. *Curr. Biol.* **20**, 25–31 (2010).
276. Bergeron, D., Pal, G., Beaulieu, Y. B., Chabot, B. & Bachand, F. Regulated Intron Retention and Nuclear Pre-mRNA Decay Contribute to PABPN1 Autoregulation. *Mol. Cell. Biol.* **35**, 2503–2517 (2015).
277. Avendaño-Vázquez, S. E. *et al.* Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev.* **26**, 1679–1684 (2012).
278. Di Gregorio, E. *et al.* A de novo X; 8 translocation creates a PTK2-THOC2 gene fusion with THOC2 expression knockdown in a patient with psychomotor retardation and congenital cerebellar hypoplasia. *J. Med. Genet.* **50**, 543–551 (2013).
279. Kumar, R. *et al.* THOC2 Mutations Implicate mRNA-Export Pathway in X-Linked Intellectual Disability. *Am. J. Hum. Genet.* **97**, 302–310 (2015).
280. Beaulieu, C. L. *et al.* Intellectual disability associated with a homozygous missense mutation in THOC6. *Orphanet J. Rare Dis.* **8**, 62 (2013).
281. Kaneb, H. M. *et al.* Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **24**, 1363–1373 (2015).
282. Nousiainen, H. O. *et al.* Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease. *Nat. Genet.* **40**, 155–157 (2008).
283. Folkmann, A. W., Noble, K. N., Cole, C. N. & Wente, S. R. Dbp5, Gle1-IP6 and Nup159: a working model for mRNP export. *Nucleus* **2**, 540–548 (2011).
284. Murphy, R., Watkins, J. L. & Wente, S. R. GLE2, a Saccharomyces cerevisiae homologue of the Schizosaccharomyces pombe export factor RAE1, is required for nuclear pore complex structure and function. *Mol. Biol. Cell* **7**, 1921–1937 (1996).
285. Bharathi, A. *et al.* The human RAE1 gene is a functional homologue of Schizosaccharomyces pombe rae1 gene involved in nuclear export of Poly(A)+ RNA. *Gene* **198**, 251–258 (1997).
286. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129–133 (2015).
287. Jovičić, A. *et al.* Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* **18**, 1226–1229 (2015).
288. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56–61 (2015).
289. Garcia-Lopez, A. *et al.* Genetic and chemical modifiers of a CUG toxicity model in Drosophila. *PLoS One* **3**, e1595 (2008).

290. Sun, X. *et al.* Nuclear retention of full-length HTT RNA is mediated by splicing factors MBNL1 and U2AF65. *Sci. Rep.* **5**, 12521 (2015).
291. Grima, J. C. *et al.* Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron* **94**, 93–107.e6 (2017).
292. D'Angelo, M. A., Raices, M., Panowski, S. H. & Hetzer, M. W. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* **136**, 284–295 (2009).
293. Mertens, J. *et al.* Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. *Cell Stem Cell* **17**, 705–718 (2015).
294. Shi, K. Y. *et al.* Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1111–E1117 (2017).
295. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* **345**, 1192–1194 (2014).
296. Eggens, V. R. *et al.* EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations. *Orphanet J. Rare Dis.* **9**, 23 (2014).
297. Rudnik-Schöneborn, S. *et al.* Pontocerebellar hypoplasia type 1: clinical spectrum and relevance of EXOSC3 mutations. *Neurology* **80**, 438–446 (2013).
298. Zanni, G. *et al.* Exome sequencing in a family with intellectual disability, early onset spasticity, and cerebellar atrophy detects a novel mutation in EXOSC3. *Neurogenetics* **14**, 247–250 (2013).
299. Boczonadi, V. *et al.* EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. *Nat. Commun.* **5**, 4287 (2014).
300. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
301. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
302. Yeo, G., Holste, D., Kreiman, G. & Burge, C. B. Variation in alternative splicing across human tissues. *Genome Biol.* **5**, R74 (2004).
303. Faustino, N. A. & Cooper, T. A. Pre-mRNA splicing and human disease. *Genes Dev.* **17**, 419–437 (2003).
304. Sassi, C. *et al.* A Novel Splice-Acceptor Site Mutation in GRN (c. 709-2 A> T) Causes Frontotemporal Dementia Spectrum in a Large Family from Southern Italy. *J. Alzheimers. Dis.* **53**, 475–485 (2016).
305. Luzzi, S. *et al.* Missense mutation in GRN gene affecting RNA splicing and plasma progranulin level in a family affected by frontotemporal lobar degeneration. *Neurobiol. Aging* **54**, 214.e1–214.e6 (2017).
306. Mukherjee, O. *et al.* Molecular characterization of novel progranulin (GRN) mutations in frontotemporal dementia. *Hum. Mutat.* **29**, 512–521 (2008).
307. Guven, G. *et al.* Mutation Frequency of the Major Frontotemporal Dementia Genes, MAPT, GRN and C9ORF72 in a Turkish Cohort of Dementia Patients. *PLoS One* **11**, e0162592 (2016).
308. Humphrey, J., Emmett, W., Fratta, P., Isaacs, A. M. & Plagnol, V. Quantitative analysis of cryptic splicing associated with TDP-43 depletion. *BMC Med. Genomics* **10**, 38 (2017).

309. Tan, Q. *et al.* Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum. Mol. Genet.* **25**, 5083–5093 (2016).
310. Jeong, Y. H. *et al.* Tdp-43 cryptic exons are highly variable between cell types. *Mol. Neurodegener.* **12**, 13 (2017).
311. Shum, E. Y. *et al.* The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. *Cell* **165**, 382–395 (2016).
312. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
313. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
314. Doucet, A. J. *et al.* Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet.* **6**, (2010).
315. Cordaux, R., Hedges, D. J., Herke, S. W. & Batzer, M. A. Estimating the retrotransposition rate of human Alu elements. *Gene* **373**, 134–137 (2006).
316. Maxwell, P. H., Burhans, W. C. & Curcio, M. J. Retrotransposition is associated with genome instability during chronological aging. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20376–20381 (2011).
317. Hua-Van, A., Le Rouzic, A., Boutin, T. S., Filée, J. & Capy, P. The struggle for life of the genome's selfish architects. *Biol. Direct* **6**, 19 (2011).
318. Ramírez, M. A. *et al.* Transcriptional and post-transcriptional regulation of retrotransposons IAP and MuERV-L affect pluripotency of mice ES cells. *Reprod. Biol. Endocrinol.* **4**, 55 (2006).
319. Aguilera, A. The connection between transcription and genomic instability. *EMBO J.* **21**, 195–201 (2002).
320. Gregersen, L. H. *et al.* MOV10 Is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. *Mol. Cell* **54**, 573–585 (2014).
321. Moldovan, J. B. & Moran, J. V. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. *PLoS Genet.* **11**, e1005121 (2015).
322. Saito, K. & Siomi, M. C. Small RNA-mediated quiescence of transposable elements in animals. *Dev. Cell* **19**, 687–697 (2010).
323. De Cecco, M. *et al.* Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging* **5**, 867–883 (2013).
324. Li, W. *et al.* Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat. Neurosci.* **16**, 529–531 (2013).
325. Douville, R., Liu, J., Rothstein, J. & Nath, A. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Ann. Neurol.* **69**, 141–151 (2011).
326. Greenwood, A. D. *et al.* Bovine spongiform encephalopathy infection alters endogenous retrovirus expression in distinct brain regions of cynomolgus macaques (Macaca fascicularis). *Mol. Neurodegener.* **6**, 44 (2011).
327. Tan, H. *et al.* Retrotransposon activation contributes to fragile X premutation rCGG-mediated neurodegeneration. *Hum. Mol. Genet.* **21**, 57–65 (2012).
328. Li, W., Jin, Y., Prazak, L., Hammell, M. & Dubnau, J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PLoS One* **7**, e44099 (2012).
329. Krug, L. *et al.* Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genet.* **13**, e1006635 (2017).

330. Subramanian, R. P., Wildschutte, J. H., Russo, C. & Coffin, J. M. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology* **8**, 90 (2011).
331. Li, W. *et al.* Human endogenous retrovirus-K contributes to motor neuron disease. *Sci. Transl. Med.* **7**, 307ra153 (2015).
332. Carroll, B., Hewitt, G. & Korolchuk, V. I. Autophagy and ageing: implications for age-related neurodegenerative diseases. *Essays Biochem.* **55**, 119–131 (2013).
333. Romano, A. D., Serviddio, G., de Matthaeis, A., Bellanti, F. & Vendemiale, G. Oxidative stress and aging. *J. Nephrol.* **23 Suppl 15**, S29–36 (2010).
334. Gensler, H. L. & Bernstein, H. DNA damage as the primary cause of aging. Q. Rev. Biol. 56, 279–303 (1981).

Figures

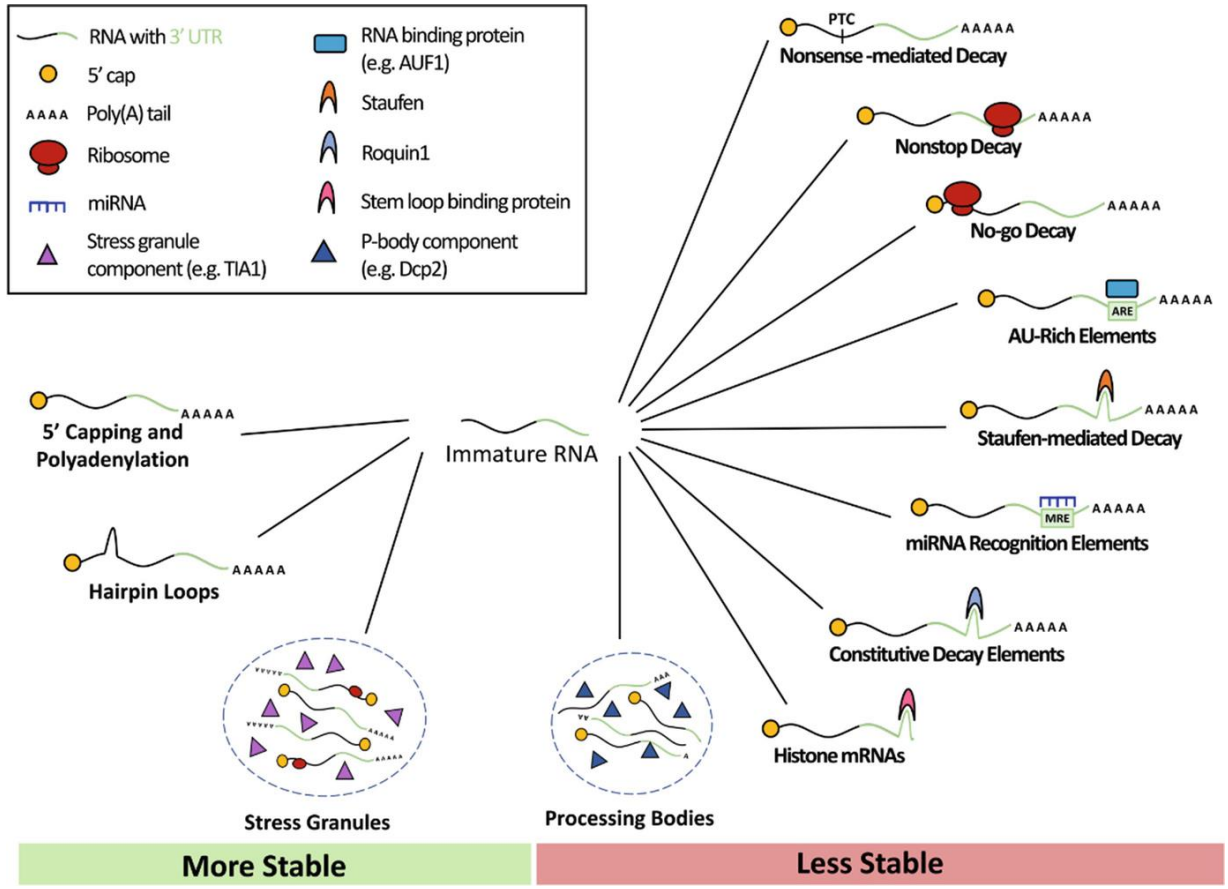


Figure 2.1. Pathways responsible for RNA homeostasis. RNA stability is promoted by two key mechanisms (left). Following transcription, nascent RNA is stabilized by the addition of a 5' cap and poly(A) tail, as well as the formation of secondary structures. Transcripts are also sequestered and stabilized in stress granules upon exposure to cellular stress. In contrast, RNA degradation pathways target faulty transcripts for removal (right). Transcripts that contain premature stop codons are targeted by nonsense-mediated decay. When translation fails to stop or start, the associated transcripts are degraded by nonstop decay and no-go decay, respectively. RNA decay mechanisms also regulate transcript abundance through several elements located within the 3' UTR, including AU-rich elements, Staufen binding sites, miRNA recognition elements, and constitutive decay elements. Lastly, P-bodies sequester and destabilize RNA transcripts.

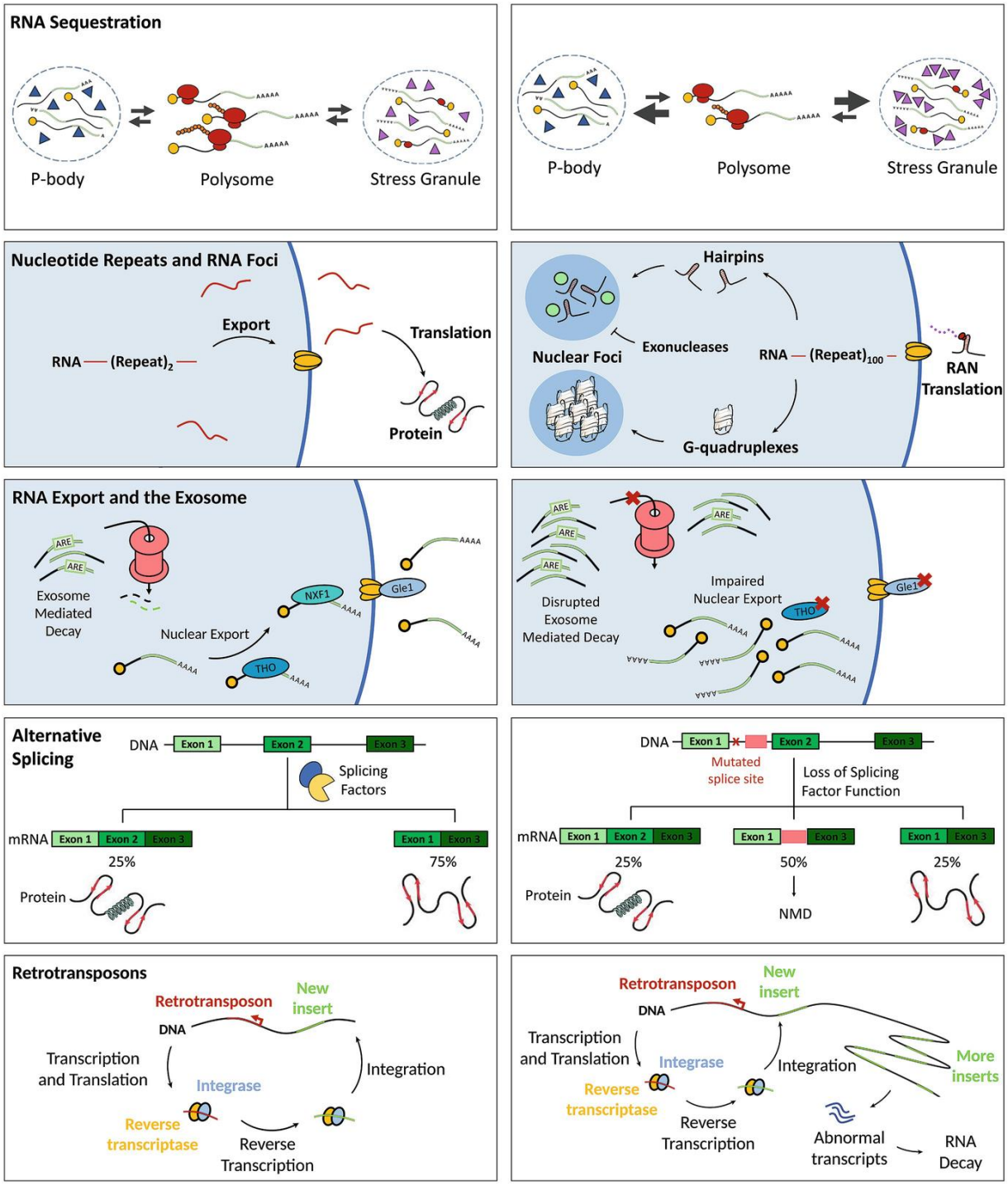


Figure 2.2. Abnormal RNA stability in neurodegenerative disease. Here we compare how normal pathways (left column) are disrupted in disease (right column). *RNA Sequestration:* There is constant flux between pools of RNA transcripts that are actively being translated (the polysome), those sequestered in stress granules, and those associated with P-bodies. In disease states, increased stress granule formation or reduced stress granule dissociation disrupts the equilibrium, resulting in fewer transcripts undergoing translation. *Repeat Expansions and RNA Foci:*

Transcripts containing repeat expansions form secondary structures such as hairpin loops and G-quadruplexes that are often stabilized in nuclear foci, which also sequester RNA-binding proteins (green circles). These transcripts also generate proteins via RAN translation that can disrupt membraneless organelles involved in RNA splicing and processing. *RNA Transport and the Exosome:* Mutations in *THO, Gle1,* and other components of the RNA export pathway result in nuclear RNA retention and degradation via the exosome complex. Mutations in exosome components can inhibit RNA turnover and further disrupt RNA homeostasis. *Alternative Splicing:* Mutations that disrupt splice sites, or dysfunction of splicing regulators such as TDP43, result in the inclusion of unannotated or "cryptic" exons (pink). These transcripts are often targeted for nonsense-mediated decay. *Retrotransposons:* These transposable elements insert themselves into the genome, often disrupting open reading frames or splice sites. The transcripts that are transcribed from these regions are often faulty, and are targeted for RNA decay.

Chapter 3. Neuronal Hyperexcitability Drives TDP43 Pathology via the Upregulation of Atypical, Shortened TDP43 Splice Isoforms*

3.1 Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder in which the progressive loss of motor neurons results in paralysis and respiratory failure[1]. There is no disease-modifying therapy for ALS, and its heterogeneous biochemical, genetic, and clinical features complicate the identification of therapeutic targets. However, the cytoplasmic mislocalization and accumulation of TDP43 (TAR DNA-binding protein of 43 kD), a nuclear RNA-binding protein integrally involved in RNA metabolism, is observed in >90% of individuals with ALS[2]. Moreover, while mutations in the gene encoding TDP43 (*TARDBP*) only account for 2-5% of ALS cases, mutations in several other ALS-associated genes including *C9orf72*[3], *ANG*[4], *TBK1*[5], *PFN1*[6], *UBQLN2*[7], *VCP*[8], and *hnRNPA2/B1*[9] result in TDP43 pathology.

TDP43 is an essential protein involved in several RNA processing events, including splicing, translation, and degradation. In keeping with these fundamental functions, TDP43 levels and localization are tightly regulated and critical for cell health. TDP43 knockout animals exhibit neurodegeneration and behavioral deficits[10–13], while TDP43 overexpression results in neurodegeneration in primary neuron[14,15], mouse[16,17], rat[18,19], Drosophila[20,21], zebrafish[22,23], and

* This chapter represents the following manuscript:

Weskamp, K. et al. Neuronal hyperexcitability drives TDP43 pathology by upregulating shortened TDP43 protein isoforms. *bioRxiv* 648477 (2019). https://doi.org/10.1101/648477

primate models[24,25]. Furthermore, mislocalization of TDP43 to the cytoplasm is sufficient to drive cell death[14]. Taken together, this suggests that even small changes to TDP43 levels and localization are highly predictive of neurodegeneration.

Hyperexcitability, or an increase in neuronal activity, is a second feature observed in both familial and sporadic ALS[26]. Cortical hyperexcitability precedes symptom onset in some cases[26], and the degree of motor neuron excitability is a strong predictor of disease progression[27,28]. Such hyperexcitability arises from a loss of cortical inhibition[26,29–33] in combination with intrinsic differences in channel expression, content, and activity within motor neurons themselves[26,28,34,35]. Emphasizing the contribution of hyperexcitability to disease, riluzole — one of two available therapies for ALS — is a sodium channel antagonist that partially rescues hyperexcitability[36]. Animal models of ALS recapitulate key features of hyperexcitability[37–39], including an increase in motor neuron activity that precedes the onset of motor deficits[37,39,40] and reduced activity following treatment with riluzole[41]. Hyperexcitability is also observed in iPSC-based ALS models[42,43], though other reports suggest that it is a transient or developmental phenomenon[43,44].

Despite the prevalence of both TDP43 pathology and hyperexcitability in ALS, the relationship between these phenomena remains poorly defined. Here, we utilize an iPSC-derived neuron (iNeuron) model system to demonstrate that hyperexcitability drives TDP43 pathology characteristic of ALS via the upregulation of atypical, shortened TDP43 isoforms. Using multiple model systems and human post-mortem material, we show that these unusual isoforms are exported from the nucleus, form insoluble cytoplasmic inclusions, are neurotoxic, and are enriched in ALS patient tissue, thereby directly implicating alternative TDP43 isoforms in ALS pathogenesis.

3.2 Results

3.2.1 TDP43 is regulated by neuronal activity

To investigate disease mechanisms related to hyperexcitability in human neurons, we established an induced pluripotent stem cell (iPSC) derived neuron (iNeuron) model. Transcription activator-like endonucleases (TALENs) specific for the *CLYBL* safe harbor locus were used to introduce the transcription factors Neurogenin 1 and 2 (Ngn1-2) under a doxycycline (dox)-inducible promoter (Figure 3.1A). Expression of Ngn1-2 is sufficient to drive the rapid differentiation of iPSCs into iNeurons that display immunocytochemical and electrophysiological properties of glutamatergic, excitatory forebrain-like neurons[45–47] (Figure 3.1B). Consistent with this, within 2 weeks of dox addition iNeurons adopt a neuronal morphology and stain positive for the neuronal markers Vglut1 and Tuj1 (Figure 3.1C). We further validated the maturity of neurons differentiated in this manner using an iPSC line that stably expresses the fluorescent calcium indicator gCaMP6f in addition to dox-inducible Ngn1-2[48]. Because time-dependent changes in gCaMP6f fluorescence correlate with action potentials, we monitored neuronal activity indirectly and non-invasively in iNeurons by fluorescence microscopy. Two to three weeks following dox addition, iNeurons displayed a low level of spontaneous activity that was significantly increased with bath application of the neurotransmitter glutamate or the potassium channel blocker tetraethylammonium (TEA; Figure 3.1D-F). Conversely, activity was inhibited by application of the sodium channel blocker tetrodotoxin (TTX). Though glutamate dramatically increased neuronal activity, it proved to be toxic even at low doses (data not shown). In comparison, iNeurons treated with TEA showed a smaller, sustained increase in activity without significant cell death (Figure 3.1G). Thus, TEA was utilized in future studies of activity-dependent TDP43 regulation.

To explore a potential connection between hyperexcitability and TDP43 pathology, we pharmacologically stimulated or blocked activity in iNeuron cultures and then examined changes in TDP43 levels via immunocytochemistry (ICC) using an antibody directed against the N-terminus. To quantitatively gauge differences in neuronal TDP43, we utilized MAP2 staining to generate cellular regions of interest (ROIs), and measured TDP43 immunoreactivity within individual neurons. TEA-treated iNeurons showed a significant increase in TDP43 immunoreactivity while TTX-treated iNeurons exhibited a reduction, suggesting a bidirectional relationship between TDP43 abundance and neuronal activity (Figure 3.1H, I). An analogous relationship was observed in rodent primary mixed cortical neurons treated with glutamate or the GABA receptor antagonist bicuculline (Supplemental Figure 3.1A). However, when we repeated these studies using an antibody directed against the TDP43 C-terminus we failed to identify significant activity-dependent changes in protein abundance (Figure 3.1J,K).

These studies also revealed prominent differences in subcellular TDP43 distribution identified by each antibody. Immunostaining with N-terminal antibodies revealed punctate, cytoplasmic TDP43 superimposed upon nuclear TDP43 in both iNeurons (Figure 3.1H) and rodent primary mixed cortical neurons treated with bicuculline (Supplemental Figure 3.1B). However, only nuclear TDP43 was detected using C-terminal TDP43 antibodies (Figure 3.1J). A survey of commercially available antibodies with known epitopes revealed similar trends in localization: antibodies that recognize the TDP43 N-terminus are more likely to display nuclear and cytoplasmic staining patterns[49–51], while antibodies specific to the C-terminus primarily show nuclear TDP43[52,53].

Given the variability in antibody specificity and potential difficulties in reproducing results using different antibodies[54,55], we validated our findings by fluorescently-labeling native TDP43

in iPSCs using CRISPR/Cas9 genome engineering. To minimize off-target effects, we used a dual-nickase approach[56] to fuse the green-fluorescent protein Dendra2 to either the N-terminus (D2-TDP43) or the C-terminus (TDP43-D2) of endogenous TDP43 in human iPSCs (Figure 3.2A, Supplemental Figure 3.2). D2-TDP43 and TDP43-D2 iPSCs were differentiated into iNeurons as described before (Figure 3.2B, C), and neuronal activity was pharmacologically stimulated or blocked using TEA or TTX, respectively. After 48h, we visualized native Dendra2-labeled TDP43 by fluorescence microscopy, noting a bidirectional relationship between D2-TDP43 abundance and neuronal activity (Figure 3.2D) that was nearly identical to what we observed using anti-TDP43 antibodies that recognize the N-terminus (Figure 3.1H, I). In comparison, there were no significant activity-dependent changes in TDP43-D2 (Figure 3.2E), consistent with our inability to detect changes upon staining with antibodies raised against the TDP43 C-terminus (Figure 3.1J, K). These data provide convincing evidence for TDP43 species harboring the N- but not the C-terminus that are regulated by neuronal activity. Additionally, the distinctive TDP43 distribution patterns revealed by N- and C-terminal reactive antibodies were reflected by the localization of Dendra2-tagged native TDP43: D2-TDP43 appeared both cytoplasmic and nuclear (Figure 3.2B), while the distribution of TDP43-D2 was limited to the nucleus (Figure 3.2C).

Collectively, these results suggest that neuronal activity elicits an increase in cytoplasmic TDP43 that lacks a C-terminus. In contrast to what we observed with N-terminal TDP43, there was no reciprocal activity-dependent change in C-terminal TDP43 abundance or localization by immunocytochemistry, and we failed to observe any differences in C-terminally labeled TDP43-D2 upon addition of TEA or TTX, arguing against a cleavage event. Previous studies demonstrated that neuronal activity regulates the abundance of similar RNA-binding proteins through alternative

splicing[57,58]. We therefore considered the possibility that activity gives rise to distinct TDP43 isoforms through alternative splicing.

3.2.2 Hyperexcitability drives *TARDBP* alternative splicing

Using available RNA-seq data obtained from human cell lines[59], we identified two alternatively spliced *TARDBP* isoforms predicted to encode C-terminally truncated or shortened (s) TDP43 isoforms (Figure 3.3A). Identical sTDP43 splice isoforms (TDP-S6 and TDP-S7) were detected in previous studies of *TARDBP* splicing[60,61]. Both sTDP43-specific splice donors are located within *TARDBP* exon 6 and differ by only 9 bp; each utilizes an identical splice acceptor within the *TARDBP* 3' untranslated region (UTR), thereby eliminating the majority of exon 6 (Figure 3.3B). We designed primers specific for both sTDP43 splice junctions as well as full-length (fl) TDP43 utilizing the canonical termination codon within exon 6, and performed qRT-PCR to examine changes in splice isoform abundance in vehicle-, TEA-, or TTX-treated human iNeurons. Both sTDP43 isoforms were not only detectable in iNeurons, but also significantly upregulated by TEA-treatment and downregulated by TTX (Figure 3.3C), suggesting that the bidirectional change in N-terminal TDP43 observed in TEA- or TTX-treated iNeurons may be due to altered expression of sTDP43 transcript isoforms. Transcripts encoding flTDP43 were also upregulated by TEA, but not reduced by TTX (Figure 3.3C). Thus, although all *TARDBP* transcript variants increase with neuronal activity, only sTDP43 isoforms demonstrate a bidirectional response to neuronal activity and corresponding changes at the protein level, perhaps due to selective autoregulation or nuclear retention of flTDP43-encoding transcripts[61–63].

The two sTDP43 transcripts (sTDP43-1 and -2) encode proteins that differ by only 3 amino acids, and both lack residues that correspond to the entirety of the glycine rich domain (residues

281-414 of flTDP43)[64]. Usage of the common splice acceptor for sTDP43-1 and -2 located within the *TARDBP* 3'UTR results in the inclusion of a new exon encoding a unique 18-amino acid C-terminus not found in flTDP43 (Figure 3.3D). These splicing events and the novel C-terminus are highly conserved at both the protein (Supplemental Table 1) and transcript (Supplemental Table 2) levels in humans, non-human primates, and lesser mammals. Despite this, and the previous identification of sTDP43 splice variants in human and murine tissues[60,61,64,65], their regulation remained unknown and unexplored. Our results demonstrate that sTDP43 variants are dynamically and bidirectionally regulated by neuronal activity, with neuronal hyperactivity resulting in a significant 2-fold upregulation of sTDP43 at the RNA and protein levels.

3.2.3 sTDP43 is cytoplasmically localized due to a putative NES in its C-terminal tail

To investigate sTDP43 localization, we transfected rodent primary mixed cortical neurons with diffusely localized mApple to enable visualization of neuronal cell bodies and processes, as well as constructs encoding flTDP43 or sTDP43-1 isoforms fused to an EGFP tag. We then imaged cultures by fluorescence microscopy to examine the localization of each isoform. As expected, flTDP43 appeared to be primarily nuclear in distribution, but sTDP43 demonstrated prominent cytoplasmic deposition (Figure 3.4A). The dramatic difference in sTDP43 localization was unanticipated given the presence of an intact nuclear localization signal (NLS) within the sTDP43 N-terminus (Figure 3.3D), and hinted at the presence of a potential nuclear export signal (NES) within the novel sTDP43 C-terminus.

To explore this possibility, we utilized NetNES1.1, an algorithm that employs neural networks and hidden Markov models to predict NES-like motifs from protein primary structure[66]. This analysis uncovered a series of amino acids near the sTDP43 C-terminal pole that could

potentially act as an NES (Figure 3.4B). We then tested the function of this putative NES through two complementary experiments. First, we altered the putative NES within sTDP43 by site-directed mutagenesis (TSLKV→GGGGG) and expressed this construct (sTDP43(mNES)) in rodent primary neurons (Figure 3.4A). Protein localization was assessed by automated microscopy, using scripts that measure fluorescence separately within cytoplasmic and nuclear ROIs, and calculate a nuclear-cytoplasmic ratio (NCR) for TDP43 in each transfected neuron[14,15]. While sTDP43 was localized to both the nucleus and cytoplasm, sTDP43(mNES) displayed a primarily nuclear distribution, more so even than flTDP43, suggesting that the putative NES is necessary for cytoplasmic deposition of sTDP43 (Figure 3.4C). Second, we fused EGFP to the 18-amino acid tail of sTDP43 (EGFP-tail), or a version of the sTDP43 tail harboring a mutated NES (EGFP-tail(mNES)) (Figure 3.4D). For comparison, we also expressed Shuttle-RFP, a construct with a strong NES and a weak NLS that exhibits a predominant cytoplasmic distribution[67]. Addition of the sTDP43 tail was sufficient to partially exclude EGFP-tail from the nucleus, but this change in distribution was eliminated by mutating the residues making up the putative NES in EGFP-tail(mNES) (Figure 3.4E). Lastly, we asked whether sTDP43's cytoplasmic distribution arises from the absence of a nuclear retention signal encoded within the canonical TDP43 C-terminus[68], or the presence of an active NES within the sTDP43 tail. Fusing the sTDP43 tail to flTDP43 markedly shifted the distribution of flTDP43 to the cytoplasm (Supplemental Figure 3.3), suggesting that sTDP43 localization is dictated primarily by the C-terminal NES, and not the lack of a nuclear retention signal. Together, these data indicate that the novel C-terminus of sTDP43 encodes a functional NES that facilitates cytoplasmic accumulation of sTDP43.

3.2.4 sTDP43 overexpression is neurotoxic

TDP43 mislocalization is a widely observed phenomenon in ALS, and cytoplasmic TDP43 is a strong predictor of cell death[14]. Given these data and the largely cytoplasmic localization of sTDP43, we surmised that sTDP43 accumulation would be toxic to mammalian neurons. We therefore utilized automated microscopy in conjunction with survival analysis to track individual neurons prospectively over time and determine their risk of death in an unbiased and high-throughput manner[14,15,59,69,70]. Rodent primary mixed cortical neurons were transfected with mApple and EGFP-tagged TDP43 isoforms and imaged by fluorescence microscopy at 24h intervals for 10d[71]. Custom scripts were used to automatically generate ROIs corresponding to each cell and determine time of death based on rounding of the soma, retraction of neurites, or loss of fluorescence (Figure 3.5A). The time of death for individual neurons was used to calculate the risk of death in each population relative to a reference group, in this case neurons expressing EGFP[71,72]. In keeping with the results of previous studies, flTDP43 overexpression resulted in a significant increase in the risk of death in comparison to EGFP alone (HR=2.22 $p<2x10^{-16}$). sTDP43-1 overexpression elicited a similar increase in the risk of death for transfected neurons (HR=1.90 $p<2x10^{-16}$), suggesting that sTDP43 and flTDP43 display similar toxicities when overexpressed in neurons (Figure 3.5B).

3.2.5 sTDP43 alters endogenous TDP43 localization

TDP43 dimerizes via its N-terminus[52,73–79], and because sTDP43 is exported from the nucleus and contains an intact N-terminus we questioned whether sTDP43 might bind to and sequester endogenous TDP43 within the cytoplasm. To determine if sTDP43 is capable of dimerizing with endogenous TDP43, we transfected HEK293T cells with HaloTag-labeled

sTDP43-1 or flTDP43 and isolated the fusion proteins using HaloLink resin (Figure 3.6A). We detected equivalent amounts of endogenous TDP43 in eluates from sTDP43-HaloTag and flTDP43-HaloTag, indicating that sTDP43 effectively binds endogenous TDP43 (Figure 3.6B).

We also examined the interaction between sTDP43 and endogenous TDP43 by ICC. HEK293T cells were transfected with EGFP or EGFP-tagged sTDP43, immunostained using a C-terminal TDP43 antibody that recognizes endogenous TDP43 but not sTDP43, and imaged by confocal fluorescence microscopy (Figure 3.6C). HEK293T cells overexpressing EGFP-tagged sTDP43 displayed cytoplasmic inclusions that strongly colocalize with endogenous TDP43. Moreover, we observed significant reductions in nuclear endogenous TDP43 in association with cytoplasmic TDP43 deposition (Figure 3.6D, E), suggesting cytoplasmic sequestration of endogenous TDP43 by sTDP43. Rodent primary mixed cortical neurons overexpressing sTDP43-1 displayed a similar depletion of endogenous TDP43 from the nucleus (Figure 3.6F, G). Thus, sTDP43 overexpression results in both cytoplasmic deposition and nuclear clearance of endogenous TDP43, recapitulating signature features of ALS pathology and implying that both gain- and loss-of-function mechanisms contribute to toxicity.

In sTDP43-transfected cells, we observed significant variability in degree of TDP43 nuclear exclusion and cytoplasmic aggregation, which we suspected was due to differences in sTDP43 expression on a per-cell basis. Because the abundance of fluorescently-labeled proteins is directly proportional to the intensity of the fluorescent tag[80], we estimated sTDP43 expression in each cell by measuring single-cell EGFP intensity, and separated cells into 5 bins based on EGFP intensity and sTDP43 expression level. Upon assessing endogenous TDP43 distribution within each bin, we observed a direct relationship between the extent of nuclear TDP43 mislocalization and sTDP43 expression (Supplemental Figure 3.4). These results indicate that TDP43 pathology

may be more prevalent with upregulation of sTDP43 via neuronal hyperexcitability or other mechanisms.

We failed to observe significant increases in cytoplasmic TDP43 deposition in transfected neurons (Figure 3.6H), potentially due to steric inhibition of sTDP43 localization and function by fusion with EGFP or HaloTag (Supplemental Figure 3.5). Placing the EGFP tag on the C-terminus of sTDP43 partially prevents cytoplasmic localization of sTDP43-EGFP (Supplemental Figure 3.5A-C), perhaps due to masking of the putative C-terminal NES. Similarly, we found that fusion of HaloTag with the N-terminus of sTDP43 significantly inhibits its binding to endogenous TDP43 (Supplemental Figure 3.5D, E). As such, N-terminal labeling of sTDP43 leaves the NES accessible but blocks association with endogenous TDP43, while C-terminal sTDP43 labeling obstructs the NES but allows interaction with endogenous TDP43.

3.2.6 sTDP43 lacks canonical functions of flTDP43

To further examine the possibility that sTDP43 elicits loss-of-function toxicity, we assessed the ability of sTDP43 to participate in TDP43-related splicing activity. In keeping with TDP43's function as a splicing repressor, TDP43 effectively blocks the inclusion of cystic fibrosis transmembrane conductance regulator (CFTR) exon 9 (Supplemental Figure 3.6A)[81,82]. In HEK293T cells expressing a CFTR minigene reporter, cotransfection with EGFP-flTDP43 resulted in proficient exon 9 exclusion as measured by PCR. In contrast, EGFP-sTDP43-1 expression failed to significantly affect exon 9 exclusion (Supplemental Figure 3.6B, C), suggesting that without the C-terminus, sTDP43 is incapable of TDP43-specific splicing[60,64,65].

Functional flTDP43 also participates in an autoregulatory feedback loop, in which flTDP43 recognizes sequences within the *TARDBP* 3'UTR, triggering alternative splicing and/or

polyadenylation and subsequent mRNA degradation[83,84]. To determine if sTDP43 is able to regulate endogenous TDP43 expression via this mechanism, we employed a TDP43 autoregulation reporter consisting of an open reading frame (ORF) encoding the fluorescent protein mCherry upstream of the *TARDBP* 3' UTR (Supplemental Figure 3.7A)[69]. In rodent primary cortical neurons expressing the TDP43 autoregulation reporter, cotransfection with EGFP-tagged flTDP43 resulted in a decrease in reporter signal, as expected. EGFP-labeled sTDP43-1 displayed more subtle effects on reporter fluorescence, suggesting that its ability to autoregulate TDP43 is impaired (Supplemental Figure 3.7B, C). Likewise, when expressed in HEK293T cells, sTDP43-1 exhibited a similarly muted effect on endogenous TDP43 itself at the transcript and protein level (Supplemental Figure 3.7D-F), consistent with poor autoregulation. Together, these results indicate that sTDP43 lacks many of the canonical functions of TDP43, including its splicing and autoregulatory abilities.

3.2.7 sTDP43 colocalizes with markers of stress granules

Previous studies suggested that sTDP43 associates with protein components of cytoplasmic stress granules, including G3BP1 and TIA1[65]. Therefore, we immunostained for G3BP1 and TIA1 in HEK293T cells overexpressing EGFP-tagged sTDP43-1, before and after application of osmotic stress (0.4M sorbitol). Prior to sorbitol treatment, we noted substantial colocalization of sTDP43-1 with G3BP1 (Supplemental Figure 3.8A) and TIA1 (data not shown) in large cytoplasmic deposits; these structures were unique to cells transfected with sTDP43-1, suggesting that sTDP43 overexpression elicits the formation of irregular structures rich in stress granule components. However, when cells were stressed with 0.4M sorbitol we observed the formation of multiple small, punctate granules that colocalize with both G3BP1 and TIA1, as well as endogenous TDP43

(Supplemental Figure 3.8B). Moreover, while osmostic stress drives flTDP43 to the cytoplasm, it has little effect on sTDP43 localization (Supplemental Figure 3.8C). These data confirm that sTDP43 localizes to stress granules, and further imply that sTDP43 production may be sufficient for the assembly of cytoplasmic stress granule-like structures even in the absence of stress.

3.2.8 sTDP43 transcripts are enriched in murine and human lumbar motor neurons

To determine if sTDP43 isoforms are produced *in vivo* and assess their expression in different regions of the central nervous system, we took advantage of a previous study that analyzed the transcriptome from murine frontal cortex and lumbar spinal motor neurons isolated by laser capture microdissection (Figure 3.7A)[85]. The most abundant splice isoform in frontal cortex homogenate was flTDP43, with predominant use of the conventional termination codon within *TARDBP* exon 6. However, splicing of the *TARDBP* locus, and in particular exon 6 and the 3'UTR, was dramatically altered in murine spinal motor neurons. In contrast to what was observed in frontal cortex, two splicing events corresponding to sTDP43 variants 1 and 2 were strongly favored in spinal motor neurons—these isoforms were upregulated ~12- and 10-fold, respectively, in lumbar spinal neurons relative to frontal cortex (Figure 3.7B, C). Further, while an ALS-associated *TARDBP* mutation (Q331K) did not affect sTDP43 transcript abundance, we noted an age-related increase in sTDP43 mRNA levels in 20-month vs. 5-month old mouse cortices (Supplemental Figure 3.9). These data show that sTDP43 isoforms are not only detectable *in vivo* within the murine CNS, but they are also significantly upregulated by age and enriched in spinal motor neurons, a cell type selectively targeted in ALS.

We also examined sTDP43 expression in human spinal neurons utilizing published RNA-seq data on laser-captured lumbar spinal motor neurons, isolated from control and sporadic (s)

ALS patient tissue[86]. Within this dataset, we identified specific transcripts corresponding to flTDP43, sTDP43-1, and sTDP43-2 based on sequence, and characterized the remaining *TARDBP* variants as "other." Although there was no apparent difference in the abundance of any *TARDBP* transcripts between sALS and control motor neurons, we noted a dramatic enrichment of sTDP43-1 transcripts in human spinal neurons, in comparison to flTDP43, sTDP43-2, and other *TARDBP* variants (Figure 3.7D, E). Furthermore, and in direct contrast to what we observed in rodent spinal neurons, sTDP43-1 was the predominant *TARDBP* splice isoform detected in human spinal neurons. To extend these findings, we also examined available RNA-seq data on spinal cord ventral horn homogenate derived from control and sALS patients[87], as well as cerebellum and frontal cortex derived from controls, sALS, and patients bearing disease-associated mutations in *C9orf72* (C9ALS)[88]. Fundamental differences in sample preparation and sequencing methodology among each study (Supplemental Figure 3.10A) prevented the direct comparison of sTDP43 abundance between tissue types. Despite these limitations, it is clear sTDP43-1 but not sTDP43-2 is expressed in several different regions of the CNS, including but not limited to spinal motor neurons, frontal cortex, and cerebellum, though it is not significantly upregulated in sALS or C9ALS (Supplemental Figure 3.10B, C).

3.2.9 Endogenously produced sTDP43 is detectable using specific antibodies

To distinguish natively-produced sTDP43 species, we generated an antibody directed against the novel 18-amino acid C-terminus of sTDP43 (Figure 3.3D). This antibody specifically recognized EGFP fused to the sTDP43 C-terminus, suggesting that the sTDP43 tail is sufficient for immunoreactivity, and the signal was completely abolished by preincubation with the immunizing peptide (Supplemental Figure 3.11A). Furthermore, expression of artificial miRNAs

(amiRNAs) targeting TDP43 effectively reduced flTDP43 levels, as expected (Supplemental Figure 3.11B, C), and also decreased sTDP43 immunoreactivity (Supplemental Figure 3.11D, E), confirming antibody specificity. We further validated the sTDP43 antibody by transfecting HEK293T cells with EGFP-tagged sTDP43-1, isolating RIPA- and urea-soluble protein fractions, and immunoblotting for sTDP43. In previous studies, overexpressed sTDP43 was highly insoluble[64]; supporting this, we detected EGFP-sTDP43 exclusively in the urea-soluble fraction while EGFP-flTDP43 appeared in both RIPA- and urea-soluble fractions (Figure 3.8A). We also tested the sTDP43 antibody in human iNeurons treated with TEA or TTX to induce or abolish neuronal activity, respectively (Figure 3.8B). In these studies, TEA increased sTDP43 immunoreactivity, while TTX reduced sTDP43 levels (Figure 3.8C), consistent with activity-dependent upregulation of N-terminally labeled D2-TDP43 (Figure 3.2) and its detection by antibodies specific for the TDP43 N-terminus (Figure 3.1). Notably, sTDP43 antibodies detected numerous cytoplasmic puncta in TEA-treated neurons that were less apparent in vehicle- and TTX-treated cells, and the background nuclear signal was minimal in all cases. Identical sTDP43-positive cytoplasmic puncta were observed in rodent primary mixed cortical neurons treated with bicuculline (Supplemental Figure 3.11F) or glutamate (data not shown). These data indicate that sTDP43 antibodies selectively detect truncated, cytoplasmic, and insoluble TDP43 species by western blot and ICC, establishing them as useful tools for investigating sTDP43 deposition and its potential role in neurodegeneration.

Based on the observed upregulation of sTDP43 splice isoforms in lumbar spinal neurons, we employed our newly-developed sTDP43 antibody for detecting sTDP43 *in vivo* within murine spinal cord sections. As predicted from the RNA-sequencing data, we detected cytoplasmic sTDP43 in anterior horn neurons from the lumbar spinal cord, confirming the subcellular

distribution of the protein originally noted *in vitro* (Supplemental Figure 3.12A). We also observed strong colocalization of sTDP43 with GFAP-positive astrocytic projections within the spinal white matter, indicating astrocytic expression of sTDP43 (Supplemental Figure 3.12B). Subsequent studies confirmed that sTDP43 is endogenously produced by human iPSC-derived astrocytes (Supplemental Figure 3.13), suggesting that while sTDP43 is enriched within spinal neurons (Figure 3.7), it is also synthesized by supporting glia.

3.2.10 sTDP43 pathology is observed in ALS patient tissue

Given that (a) sTDP43 is endogenously produced at relatively high levels in spinal neurons, (b) neuronal hyperexcitability is a conserved feature of ALS, and (c) sTDP43 is upregulated by neuronal activity, we suspected that sTDP43 may accumulate in individuals with sALS. To address this question, we immunostained human cortex and spinal cord sections from sALS, C9ALS, and unaffected control patients using antibodies that recognize the TDP43 N-terminus or our newly-developed sTDP43 antibodies (Figure 3.8D). As predicted, immunostaining with N-terminal TDP43 antibodies showed both a reduction in nuclear signal and the appearance of cytoplasmic inclusions selectively in ALS patient tissue. While control tissue exhibited low immunoreactivity for sTDP43 in both the cortex and spinal cord, we observed a striking accumulation of sTDP43 within cytoplasmic deposits in ALS spinal cord and cortex (Figure 3.8E). sTDP43-positive inclusions closely colocalized with N-terminally reactive cytoplasmic aggregates but not residual nuclear TDP43, suggesting that sTDP43 antibodies specifically label cytoplasmic deposits in ALS tissue. In this limited case study, sTDP43 pathology appeared to be conserved between sALS and C9ALS, hinting at a conserved process (Supplemental Figure 3.14). We also observed a tight correlation between conventional TDP43 pathology (nuclear exclusion and cytoplasmic

aggregation) and sTDP43 deposition—in ALS samples, neurons displaying TDP43 nuclear exclusion almost always showed cytoplasmic sTDP43 pathology. Additionally, we noted several cells from ALS patients that exhibited sTDP43 deposits despite a normal nuclear TDP43 pattern, perhaps illustrating an early stage of pathology (Supplemental Figure 3.14B, C). Even so, we detected significant heterogeneity in sTDP43 pathology among ALS cases, indicating the presence of additional, unknown factors that could impact sTDP43 deposition or immunoreactivity.

In light of endogenous sTDP43 detected within mouse spinal cord astrocytes (Supplemental Figure 3.12B) and human iPSC-derived astrocytes (Supplemental Figure 3.13), we asked if sTDP43 pathology might also be present within astrocytes. In sections from controls and sALS patients immunostained with sTDP43 antibodies, neurons and glia were identified by co-staining with NeuN and GFAP antibodies, respectively (Figure 3.8F). Cytoplasmic sTDP43 accumulations were detected in both NeuN-positive neurons and GFAP-positive astrocytes, suggesting that sTDP43 pathology is not limited to neurons. Taken together, these results demonstrate that endogenous sTDP43 accumulates within neurons and glia of individuals with ALS, supporting a potentially pathogenic contribution of sTDP43 isoforms to ALS pathogenesis.

3.3 Discussion

In this study, we show that neuronal hyperactivity leads to the selective upregulation of C-terminally truncated TDP43 isoforms (sTDP43). These isoforms are intrinsically insoluble and accumulate within cytoplasmic aggregates by virtue of a NES present within a novel 18-amino acid C-terminus. sTDP43 also sequesters endogenous TDP43 within cytoplasmic aggregates and induces its clearance from the nucleus, thereby recapitulating signature pathologic changes found in the majority of individuals with ALS and implicating complementary gain- and loss-of-function

mechanisms in disease pathogenesis. sTDP43 transcripts are enriched in murine motor neurons, a cell type that is selectively vulnerable in ALS, and post-mortem samples from individuals with ALS show conspicuous accumulations of sTDP43 within affected neurons and glia. These observations suggest a fundamental link between neuronal hyperexcitability and TDP43 pathology, two conserved features characteristic of both familial and sporadic ALS. Moreover, they raise the possibility that sTDP43 production and/or its accumulation are heretofore-unrecognized contributors to neurodegeneration in ALS.

A series of previous studies demonstrated that alternative splicing of *TARDBP* gives rise to truncated TDP43 isoforms lacking the C-terminus that are highly insoluble when overexpressed in heterologous systems[49,60,64]. Here, we show for the first time that neuronal activity selectively upregulates these truncated isoforms, which we collectively labeled sTDP43, despite a simultaneous increase in transcripts encoding full-length TDP43. This discrepancy may arise from the relative inability of sTDP43 to effectively participate in autoregulation (Supplemental Figure 3.7), or the presence of unique elements within the flTDP43 3'UTR leading to nuclear mRNA retention and/or destabilization[61–63,83,84]. As such, the activity-dependent and apparently selective upregulation of sTDP43, together with the widespread neuronal hyperactivity observed in ALS patients, animal models, and human iPSC-derived neurons[26,28,34,35,42,43], may be a crucial factor driving sTDP43 deposition in ALS tissue.

In keeping with previous studies[65], overexpressed sTDP43 accumulates in the cytoplasm where it often forms large, insoluble inclusions. The low-complexity domain (LCD) within the TDP43 C-terminus is essential for liquid-phase separation[89–92], and has been heavily implicated in TDP43 aggregation. Even so, our observations and those of others[64,65] show that sTDP43 is insoluble and prone to aggregation, despite lacking the LCD. A growing body of evidence suggests

that proteins with complex, folded domains such as the TDP43 RNA-recognition motifs (RRMs) are highly susceptible to aggregation[93]. Rather than promoting insolubility, the presence of LCDs within these proteins protects against misfolding and aggregation by enabling reversible phase transitions during conditions of supersaturation. Thus, LCDs may permit higher local concentrations of RRM-containing proteins than would otherwise be possible without misfolding and/or aggregation[94]. In this regard, the absence of the LCD may be directly responsible for the enhanced aggregation of sTDP43; indeed, several RNA-binding proteins display similar phenotypes upon removal of the LCD, including Pub1, Pab1 and Sup35[94–97].

Using predictive software, we identified a potential NES located within the novel 18-amino acid sTDP43 tail, and experimentally confirmed that this segment drives cytoplasmic sTDP43 localization. This NES appears to be dominant over the functional NLS present within the N-terminus of sTDP43, either due to a high affinity for nuclear exporters or because of enhanced accessibility of the NES at the extreme C-terminus of the protein. The previously annotated TDP43 NES[68,98] exhibits leucine/isoleucine-rich sequences favored by exportin-1 (XPO1), an essential mediator of protein nuclear export[99]. Nevertheless, scant experimental evidence suggests that this sequence functions as a true NES. TDP43 and XPO1 do not interact with one another *in vitro*[15,100], and unbiased proteomics studies have failed to identify TDP43 as an XPO1 cargo protein[101,102]. Further, TDP43 localization is unaffected by XPO1 inhibition or deletion of the putative NES[15]. In contrast, the NES uncovered within the sTDP43 C-terminal tail is both necessary and sufficient for sTDP43 nuclear export, suggesting that it is a *bona fide* NES.

sTDP43 lacks the C-terminal glycine rich domain required for splicing activity[103]; as such, sTDP43 is incapable of *CFTR* minigene splicing or effectively participating in TDP43 autoregulation, which involves differential splicing of the *TARDBP* 3'UTR[83,84]. The C-terminal

glycine rich domain is also required for toxicity upon TDP43 overexpression in yeast[104]. Nevertheless, sTDP43 overexpression was still lethal in neurons. We suspect that sTDP43-related toxicity arises from a combination of factors, including (a) the NES within the new C-terminal tail region provoking cytoplasmic sTDP43 deposition; (b) its interaction with endogenous flTDP43 via its N-terminus[52,73,105]; and (c) the presence of intact RNA-recognition motifs that enable sTDP43 to bind and potentially sequester cytoplasmic mRNAs.

sTDP43 isoforms are highly conserved in humans, non-human primates, and lesser mammals at the transcript and protein levels. Such evolutionary conservation suggests that these isoforms fulfill unknown functions, perhaps involving a compensatory response to chronic neuronal hyperactivity or generalized stress. Intriguingly, sTDP43 transcripts are significantly enriched in murine motor neurons compared to frontal cortex homogenate, their expression increases with age, and sTDP43-1 is the dominant *TARDBP* species in human lumbar motor neurons, raising the possibility that spinal motor neurons accumulate potentially toxic levels of sTDP43 in response to aging and hyperexcitability. Future studies are needed to determine whether native sTDP43 performs an essential function in motor neurons or other cell types, and if sTDP43 contributes to the selective vulnerability of motor neurons in ALS[106,107].

By creating a unique antibody that recognizes the novel C-terminus of sTDP43, we detected cytoplasmic sTDP43 inclusions selectively within the spinal cord and cortex of ALS patients, including individuals with sALS and C9ALS. In addition, the presence of sTDP43 deposits coincided with nuclear TDP43 exclusion, as predicted by sTDP43 nuclear export and its affinity for flTDP43. Although the aggregation-prone[108] TDP43 C-terminus forms a core component of the cytoplasmic inclusions found in ALS patients[109–116], emerging evidence suggests that N-terminal TDP43 fragments also contribute to ALS pathogenesis. N-terminal TDP43

fragments are observed in ALS patient spinal cord[117,118], and in keeping with studies of RNA-binding proteins in yeast, the TDP43 RRMs misfold and aggregate in vitro without the C-terminal LCD to maintain solubility[65,93,95–97,119–121]. Independent of the RRMs, the TDP43 N-terminus enhances TDP43 aggregation and toxicity[65,79,119,120], potentially adding to sTDP43 insolubility and the impact of sTDP43 deposition in affected neurons.

TDP43-positive cytoplasmic inclusions in ALS are not limited to neurons but are also found in astrocytes and oligodendrocytes[122–125]. Astrocytes help regulate extracellular glutamate levels, and their dysfunction in ALS may lead to impaired synaptic glutamate buffering in sporadic as well as familial ALS[126–131]. In addition to detecting endogenous sTDP43 production in cultured human astrocytes and murine spinal cord, we noted disease-specific astrocyte sTDP43 pathology in sALS patient tissue. Although the effect of sTDP43 accumulation in these cells remains to be determined, it is possible that sTDP43-induced astrocyte toxicity triggers a feed-forward mechanism in which reduced glutamate buffering results in neuronal hyperactivity, increased sTDP43 production, and subsequent neurodegeneration.

Our work underlines the significance of previously identified TDP43 isoforms and highlights a pivotal connection between neuronal hyperexcitability and TDP43 pathology, two conserved findings in ALS. Many questions remain, including the function of sTDP43 isoforms, the extent and pervasiveness of sTDP43 pathology in ALS, and whether cell type- or species-specific differences in sTDP43 expression contribute to the selective vulnerability of human motor neurons in ALS. Complementary investigations of sTDP43 splicing and its regulation are crucial if we are to determine if targeted manipulation of sTDP43 has the potential to prevent or slow motor neuron degeneration in ALS.

3.4 Materials and Methods

3.4.1 Generation and maintenance of iPSCs

Fibroblasts were reprogrammed into iPSCs via transfection with episomal vectors encoding seven reprogramming factors[132] and validated as previously described[133]. All iPSC lines were cultured in Essential 8 (E8) media (Gibco A1517001) on plates coated with vitronectin (Gibco A14700) diluted 1:100 in Mg^{2+}/Ca^{2+}-free phosphate buffered saline (PBS, Gibco 14190-144). Cells were passaged every 5–6d using 0.5 mM EDTA (Sigma E7889) dissolved in PBS followed by gentle trituration in E8 media with a P1000 pipette. All lines are verified mycoplasma-free on a monthly basis.

3.4.2 Integration of Ngn1/Ngn2 cassette into iPSCs

iPSCs were split and plated into a vitronectin-coated 6 well plate as described above, at a density such that cells were 50-70% confluent in clumps of 2-5 cells at the time of transfection. Following plating, cells were incubated overnight in E8 media with ROCK inhibitor (Fisher BDB562822), and changed into fresh E8 media the following morning. Thirty minutes prior to transfection (~24h after plating or when the density was 50-70%), cells were changed into mTESR-1 media (Cell Technologies 85850) and then transfected with 2.5 µg of donor DNA and 1.25 µg of each targeting construct (Supplemental Table 3) using Lipofectamine Stem (Invitrogen STEM00003) according to the manufacturer's instructions. The following morning, cells were changed into fresh E8 media. Media was changed daily, and cells were screened for red fluorescence. When the partially positive colonies reached 100-500 cells, they were carefully scraped/aspirated using a P200 pipet tip and transferred to a new vitronectin-coated dish. This process was repeated, enriching the fluorescent cells until a 100% fluorescent colony was identified. This was then relocated to a new dish, and

expanded for future use. The Ngn1/2 integration cassette and accompanying targeting constructs were a gift from M. Ward.

3.4.3 iNeuron differentiation

Day 0. Induced pluripotent stem cells were washed in PBS and incubated in prewarmed accutase (Sigma A6964) at 37°C for 8m. Four volumes of E8 media were added to the plate, and the cells were collected and pelleted at 200xg for 5m. The media was aspirated, and the pellet was resuspended in 1ml of fresh E8 media. Cells were counted using a hemocytometer, diluted, plated at a density of 20,000 cells/ml in E8 media with ROCK inhibitor and incubated at 37°C overnight. **Day 1.** Media was changed to N2 media (1x N2 Supplement (Gibco 17502-048), 1x NEAA Supplement (Gibco 11140-050), 10 ng/ml BDNF (Peprotech 450-02), 10 ng/ml NT3 (Peprotech 450-03), 0.2 µg/ml laminin (Sigma L2020), 2 mg/ml doxycycline (Sigma D3447) in E8 media). **Day 2.** Media was changed to transition media ((1x N2 Supplement, 1x NEAA Supplement, 10 ng/ml BDNF, 10 ng/ml NT3, 0.2 µg/ml laminin, 2 mg/ml doxycycline in half E8 media, half DMEM F12 (Gibco 11320-033)). **Day 3.** Media was changed into B27 media (1x B27 Supplement (Gibco 17504-044), 1x Glutamax Supplement (Gibco 35050-061), 10 ng/ml BDNF, 10 ng/ml NT3, 0.2 µg/ml laminin, 2 mg/ml doxycycline, and 1x Culture One (Gibco A33202-01) in Neurobasal-A (Gibco 12349-015)). **Day 6.** An equal volume of B27 media without Culture One was added to each well. **Day 9-21.** All cultures underwent a half-media change every 3d in fresh B27 media.

3.4.4 Immunocytochemistry

Neurons were fixed with 4% paraformaldehyde (PFA; Sigma P6148), rinsed with PBS, and permeabilized with 0.1% Triton X-100 (Bio-rad 161-0407) in PBS. Neurons were then treated with 10 mM glycine (Fisher BP381-1) in PBS, and incubated in a blocking solution (0.1% Triton X-100, 2% fetal calf serum (Sigma F4135), and 3% bovine serum albumin (BSA, Fisher BP9703-100) in PBS) at room temperature for 1h before incubation overnight at 4°C in primary antibody diluted in blocking buffer (Supplemental Table 4). Cells were then washed 3x in PBS and incubated at room temperature with Alexa Fluor 488 goat anti-rabbit (Life Technologies A11034), Alexa Fluor goat anti-mouse 594 (Life Technologies A11032), or Alexa Fluor donkey anti-rabbit 647 (Life Technologies A31573) secondary antibody diluted 1:5000 in blocking solution for 1h. Following 3x washes in PBS containing 1:10000 Hoechst 33258 dye (Invitrogen H3569), neurons were imaged via fluorescence microscopy. High resolution images were obtained on a Zeiss LSM 800 with a 63x NA1.4 Oil/DIS Plan-Apochromat objective. Excitation was accomplished using 405, 488, 561, and 633 nm lasers.

3.4.5 Modulation of neuronal activity

Half of the existing media was removed from rodent cortical neurons or mature human iNeurons and replaced with fresh media and drug such that the final concentration on the cells was 4 mM tetraethylammonium chloride (TEA, Sigma T2265)**,** 2 µM tetrodotoxin citrate (TTX, R&D Systems 1078) or 25 mM glutamate (Sigma G1251) alongside a volume-matched vehicle control. Cells were incubated at 37° C for 48h, then fixed, imaged, or harvested as needed.

Monitoring calcium transients

Mature iNeurons—differentiated as previously described from an iPSC line stably expressing gCaMP6f and mCherry[48]—were imaged for 100ms at 200ms intervals for a total of 100 frames, for a cumulative a 20s observation window. One location was imaged per well for 2-30 instances over a 6-12h period. Each neuron was identified as a region of interest using mCherry fluorescence, and the intensity of gCaMP6f signal was plotted over time. Individual traces were corrected for photobleaching, normalized to the median of each imaging period, and filtered for peaks below a discrete threshold to aid in spike identification. The number of peaks for each neuron and each imaging period was manually counted using a custom-designed graphical user interface. Events per second were averaged for each cell and compared across groups.

3.4.6 CRISPR/Cas9 editing of iPSCs

Oligos complementary to the target region (Supplemental Table 3) were annealed, digested, and ligated into the *BbsI* site in pX335-U6-Chimeric_BB-CBh-hSpCas9n(D10A) (Addgene #42335, deposited by Feng Zhang) or pX330S-4 (Addgene #58780, deposited by Feng Zhang) according to the protocol available from Addgene. iPSCs stably expressing Ngn1-2 under a dox-inducible promoter were split and transfected as described above with pX335 vectors encoding Cas9(D10A) and sgRNA pairs targeting sequences flanking the *TARDBP* start codon for D2-TDP43 or stop codon for TDP43-D2. Cells were cotransfected with the appropriate HDR vector encoding the Dendra2 open reading frame flanked by 400 bp of sequence homologous to that surrounding the *TARDBP* start codon (D2-TDP43) or stop codon (TDP43-D2) (in pUC-minus(M), synthesized by Blue Heron, LLC). Fluorescent cells were selected and successively passaged as described above to generate iPSC colonies in which 100% of cells expressed Dendra2-labeled TDP43.

3.4.7 Longitudinal fluorescence microscopy and automated image analysis

Neurons were imaged as described previously[69,70] using a Nikon Eclipse Ti inverted microscope with PerfectFocus3a 20X objective lens and either an Andor iXon3 897 EMCCD camera or Andor Zyla4.2 (+) sCMOS camera. A Lambda XL Xenon lamp (Sutter) with 5 mm liquid light guide (Sutter) was used to illuminate samples, and custom scripts written in Beanshell for use in μManager controlled all stage movements, shutters, and filters. Custom ImageJ/Fiji macros and Python scripts were used to identify neurons and draw both cellular and nuclear regions of interest (ROIs) based upon size, morphology, and fluorescence intensity. Fluorescence intensity of labeled proteins was used to determine protein localization or abundance. Custom Python scripts were used to track ROIs over time, and cell death marked a set of criteria that include rounding of the soma, loss of fluorescence and degeneration of neuritic processes[71].

3.4.8 RNA sequencing

Raw reads from murine frontal cortex and spinal cord[59,85,86], as well as human spinal cord, frontal cortex and cerebellum[59,85,86], were downloaded from Gene Expression Omnibus (GEO) with the SRA Toolkit v2.9.2. Reads were trimmed with TrimGalore v0.6.0 using automatic adapter detection and a minimum Phred score of 20. For alignment-free transcript-level quantification, trimmed reads were quantified using Salmon v0.13.1 (Patro 2017) and imported into the RStudio using txImport v1.12.0 (Soneson 2015) to generate transcript-level summaries[134,135]. The Ensembl genome assemblies and transcript annotations from GRCh38.96 and GRcm38.96 were used as human and mouse references, respectively. For alignment-based analysis of mouse datasets, trimmed reads were aligned with hisat2 v2.0.5 and raw counts were quantified for unique splice

donor/acceptor combinations present in unique *TARDBP* isoforms. In each case, splicing events were visualized using IG Viewer software (Broad Institute).

3.4.9 qRT-PCR

RNA was isolated using the RNeasy Mini Kit (Qiagen 74106), and cDNA was reverse transcribed from 1 ug of the resultant RNA with the Bio-Rad iScript kit (Bio-Rad 170-8891) in a reaction volume of 20 µl. 0.5 µl of cDNA was used for each reaction as a template for quantitative (q)PCR, which was performed using Power SYBR Green (Applied Biosystems A25742) using the primers listed in Supplemental Table 5.

3.4.10 Plasmids

Plasmids pGW1-EGFP(1)[80], pGW1-TDP43-EGFP[136], and pGW1-mApple[136] were used both as experimental controls and to generate additional constructs (Supplemental Table 6).

To generate pGW1-sTDP43-EGFP, a geneblock comprised of the sTPD43-1 open reading frame (ORF) flanked by ApaI and AgeI restriction enzyme sites was generated by Integrated DNA Technologies (IDT). This geneblock was digested with ApaI and AgeI and cloned into the corresponding sites immediately upstream of the EGFP ORF in pGW1-EGFP(1).

To create pGW1-sTDP43(mNES)-EGFP, the sTDP43 open reading frame was amplified using a reverse primer to mutate the putative NES into five sequential glycine resides. The resulting amplicon was digested with ApaI and AgeI and cloned into corresponding sites in pGW1-EGFP(1).

To generate pGW1-EGFP(2), the EGFP open reading frame was PCR amplified from pGW1-EGFP(1). The resultant amplicon was digested with HindIII and Kpn1 restriction enzymes and cloned into the corresponding sites in pGW1-CMV.

To generate pGW1-EGFP-tail, sense and antisense oligomers with the sequence of the 18-amino acid tail were generated by IDT, designed such that annealing would result in cohesive ends identical to cut KpnI and NheI restriction enzyme sites. The annealed oligo was cloned into corresponding sites immediately downstream of the EGFP ORF in pGW1-EGFP(2).

To generate pGW1-EGFP-tail(mNES), sense and antisense oligomers with the sequence of the 18-amino acid tail in which the putative NES was replaced by 5 glycine residues were generated by IDT, designed such that annealing would result in cohesive ends identical to cut KpnI and NheI restriction enzyme sites. The annealed oligo was cloned into corresponding sites immediately downstream of the EGFP ORF in pGW1-EGFP(2).

To create pGW1-EGFP-TDP43, the TDP43 ORF was PCR amplified from pGW1-TDP43-EGFP. The resultant amplicon was digested with KpnI and NheI restriction enzymes and cloned into the corresponding sites immediately downstream of the EGFP ORF in pGW1-EGFP(2).

To generate pGW1-EGFP-sTDP43, the sTDP43-1 ORF was PCR amplified from pGW1-sTDP43-EGFP. The resultant amplicon was digested with KpnI and NheI restriction enzymes and cloned into the corresponding sites immediately downstream of the EGFP ORF in pGW1-EGFP(2).

To create pGW1-Halo, the HaloTag ORF was PCR amplified from pFN21A HaloTag PUM2 RBD R6SYE (a gift from A. Goldstrohm). The resultant amplicon was digested with XbaI and SbfI restriction enzymes and cloned into the corresponding sites in pGW1-CMV.

To create pGW1-TDP43-Halo, the TDP43 ORF was PCR amplified from pGW1-TDP43-EGFP. The resultant amplicon was digested with NheI and AgeI and cloned into corresponding sites in pGW1 to make pGW1-TDP43. The HaloTag ORF was then amplified from pFN21A HaloTag PUM2 RBD R6SYE, digested with XbaI and SbfI restriction enzyme sites and cloned into the corresponding sites immediately downstream of the TDP43 ORF in pGW1-TDP43.

To generate pGW1-sTDP43-Halo, the sTDP43-1 ORF was PCR amplified from pGW1-sTDP43-EGFP. The resultant amplicon was digested with AgeI and NheI and cloned into corresponding sites immediately upstream of the HaloTag ORF in pGW1-Halo.

To generate the TDP43 autoregulatory reporter[136], a 3 kb segment extending from *TARDBP* exon 6 to the 3' UTR was amplified from genomic DNA. The resultant amplicon was digested with BsrGI and SfiI and cloned into corresponding sites immediately downstream of the mCherry ORF in pCAGGs-mCherry.

Shuttle-RFP (pcDNA3.1-NLS-mCherry-NES) was purchased from Addgene (#72660, donated by B. Di Ventura and R. Eils). The CFTR minigene reporter was a gift from Y. Ayala, and pCaggs-

mCherry and pGW1-CMV were gifts from S. Finkbeiner. The shTARDBP and non-targeting shRNA constructs were purchased from Dharmacon (V3SH11240-224779127, VSC11712).

All constructs were verified by Sanger sequencing, and described in Supplemental Table 6.

3.4.10 Primary neuron cell culture and transfection

Cortices from embryonic day (E)19-20 Long-Evans rat embryos were dissected and disassociated, and primary neurons were plated at a density of $6x10^5$ cells/ml in 96-well plates, as described previously[137]. At *in vitro* day (DIV) 4, neurons were transfected with 100 ng pGW1-mApple to mark cell bodies and 50-100 ng of an experimental construct using Lipofectamine 2000 (Invitrogen 52887), as previously described[14,71,136]. Following transfection, cells were placed in either Neurobasal Complete Media (Neurobasal (Gibco 21103-049), 1x B27, 1x Glutamax, 100 units/mL Pen Strep (Gibco 15140-122)) or NEUMO photostable medium with SOS (Cell Guidance Systems M07-500) and incubated at 37°C in 5% CO_2.

3.4.11 Culturing and transfecting HEK293Ts

Human embryonic kidney (HEK) 293T cells were cultured in DMEM (GIBCO 11995065), 10% FBS (Gibco ILT10082147), 1x Glutamax, and 100 units/mL Pen Strep at 37°C in 5% CO_2. HEK293T cells are originally female in origin, are easily transfected, and have been transformed with SV40 T-antigen. HEK293T cells were transfected with Lipofectamine 2000 according to the manufacturer's instructions.

3.4.12 Immunoprecipitation using HaloLink

HEK293T cells were transfected with HaloTagged constructs of interest. Two days after transfection, cells were collected in PBS and pelleted at 21,000xg for 5m. The cells were then resuspended in 100 µl lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate). After incubation on ice for 15m, cells were passed through a 27.5 G needle and pelleted at 21,000xg for 10m at 4° C. 100 µg of protein was then added to 100 µl of prewashed HaloLink resin (Promega G1914), which was prepared by washing and pelleting for 2m at 800xg 3x in wash buffer (100 mM Tris pH 7.5, 150 mM NaCl, 1 mg/ml BSA, 0.005% IGPAL). Sufficient wash buffer was added to ensure an equal volume for all conditions (~400 µl), and samples were incubated on a tube rotator for 30m at room temperature. Samples were then pelleted at 800xg for 2m, saving the supernatant. The beads were then washed 3x in wash buffer, and resuspended in elution buffer (1% SDS, 50 mM Tris-HCl pH 7.5) and 10x sample buffer (10% SDS, 20% glycerol, 0.0025% bromophenol blue, 100 mM EDTA, 1 M DTT, 20 mM Tris, pH 8). Samples were then boiled at 95° C for 10m, and loaded onto a 10% SDS-PAGE gel alongside 10 µg of input protein and a fixed volume of supernatant to assess binding efficiency. The gel was run at 120 V, and samples were then transferred at 100 V at 4°C onto an activated 2 um polyvinylidene difluoride (PVDF) membrane (Bio-Rad 1620177), blocked with 3% BSA in 0.2% Tween-20 (Sigma P9614) in Tris-buffered saline (TBST) for 1h, and blotted overnight at 4°C with primary antibody in 3% BSA in TBST (Supplemental Table 4). The following day, blots were washed in 3x in TBST, incubated at room temperature for 1h with donkey anti-mouse 680 RD (Li-Cor 926-68072) and donkey anti-rabbit 800 CW (Li-Cor 925-32213) secondary antibodies, both diluted 1:5,000 in 3% BSA in TBST. Following treatment with secondary antibody, blots were washed 3x in TBST, placed in Tris-buffered saline, and imaged using an Odyssey CLx Imaging System (LI-COR).

3.4.13 Differential solubility fractionation

HEK293T cells were transfected in a 6-well plate with 3 µg of DNA/well using Lipofectamine 2000 according to the manufacturer's instructions. Two days after transfection, cells were collected in PBS and pelleted at 21,000xg for 5m. Cells were then resuspended in RIPA buffer (Thermo Scientific 89900) with protease inhibitors (Roche 11836170001) and incubated on ice for 15m. Lysates were then sonicated at 80% amplitude with 5s on/5s off for a total of 2m using a Fisher Brand Model 505 Sonic Dismembrenator (ThermoFisher). Samples were centrifuged at 21,000xg for 15m at 4°C, after which the supernatant was removed and saved as the RIPA-soluble fraction. The RIPA-insoluble pellet was washed in RIPA once more and resuspended in urea buffer (7 M urea, 2 M thiourea, 4% CHAPS, 30 mM Tris, pH 8.5) and incubated on ice for 5m. Samples were then centrifuged at 21,000xg for 15m at 4°C, and the supernatant was saved as the RIPA-insoluble, urea-soluble fraction. The RIPA-soluble samples were quantified and 10-30 µg of protein/well was diluted in RIPA buffer with 10x sample buffer. For urea fractions, equal volumes of each sample across conditions was diluted in urea buffer and 10x sample buffer. The RIPA-soluble samples were boiled for 10m before 10-30 µg of all samples were loaded onto a 10% SDS-PAGE gel with stacking gel and run at 120 V. The blot was then transferred and probed as described above.

3.4.14 Immunohistochemistry in human tissue

Paraffin-embedded human cortex and spinal cord obtained from the University of Michigan Brain Bank were cut into 5 µm thick sections and mounted on glass slides. Tissue samples were photobleached prior to immunofluorescence[138]. Briefly, slides were placed on ice, under a 7 Band

Spectrum LED Light (HTG Supply 432W HTG-432-3W-7X) at 4°C for 12h. Slides were deparaffinized at 65°C for 20m, and rehydrated 5m sequentially in xylene (Fisher X3S-4), 100% ethanol (Fisher 3.8L), 95% ethanol, 70% ethanol, 50% ethanol, and PBS. Slides were then permeabilized with 0.1% Triton X-100 in PBS, and treated with 10 mM glycine in PBS. They were then incubated in a blocking solution (0.1% Triton X-100, 2% fetal calf serum, and 3% BSA in PBS) at room temperature for 1h before incubation overnight at 4°C in primary antibody diluted in blocking buffer (Supplemental Table 4). Slides were then washed 3x in PBS and incubated at room temperature with Alexa Fluor 488 goat anti-rabbit (Life Technologies A11034), Alexa Fluor goat anti-mouse 594 (Life Technologies A11032), and/or Alexa Fluor goat anti-chicken 647 (Life Technologies A21449) secondary antibody diluted 1:5000 in blocking solution for 1h. Following 3x washes in PBS containing 1:10,000 Hoechst 33258 dye (Invitrogen H3569), slides were mounted in mounting media (Fisher SP15-500) and allowed to dry in the dark overnight before being imaged the following day. Images were acquired using a Nikon Microphot-FXA microscope (Nikon, 1985) in combination with a 60x oil-immersion objective, a QIClick CCD Camera (Q Imaging, 7400-82-A1), and an X-Cite Series 120 light source (Lumen Dynamics).

3.4.15 Statistical analysis

Statistical analyses were performed in R or Graphpad Prism 7. For primary neuron survival analysis, the open-source R survival package was used to determine hazard ratios describing the relative survival between conditions through Cox proportional hazards analysis[71]. Significance determined via the two-tailed t-test was used to assess differences between treatment groups for neuronal activity and transcript abundance via RT-PCR. The Kolmogorov-Smirnov test was used to assess differences between the distribution of TDP43 abundance in neurons under different

activity conditions. One-way ANOVA with Tukey's or Dunnett's post-tests were used to assess significant differences among nuclear/cytoplasmic ratios, binding affinity, TDP43 splicing activity, and TDP43 autoregulation. Data are shown as mean ± SEM unless otherwise stated.

3.5 Supplemental Materials and Methods

3.5.1 RNA-sequencing

We utilized previously described RNA sequencing(seq) data[139,140], and determined transcript abundance as described above.

3.5.2 TDP43 knockdown in N2A cells

N2A mouse neuroblastoma cells were cultured in DMEM (GIBCO 11995065), 10% FBS (Gibco ILT10082147), 1x Glutamax, and 100 units/mL penicillin/streptomycin at 37°C in 5% CO_2. Cells were transfected using Lipofectamine 2000 (ThermoFisher) according to the manufacturer's instructions, with artificial microRNAs (amiRNAs) directed against *TARDBP* or a scrambled control[141]. Transfected N2As were incubated for 72h, harvested, and immunoblotted to verify TDP43 knockdown.

3.5.3 Tissue preparation and immunohistochemistry in murine tissue

Vertebral columns were dissected from 5 month old C57Bl6 J mice, fixed in 4% paraformaldehyde (PFA) at 4°C for 48h, washed in phosphate buffered saline (PBS), and dissected to extract the spinal cords. The lumbar enlargement was sub-dissected, cryoprotected in 30% sucrose at 4°C, embedded in M1 matrix (Thermo Scientific #1310) in a silicone mold, frozen on dry ice, and sectioned at a thickness of 16μm onto charged slides (Thermo Scientific J1800AMNZ). Sections

were then briefly air dried and stored at -80°C. For immunohistochemistry (IHC), sections were washed in distilled water, and blocked and permeabilized in blocking buffer (5% BSA, 0.5% Triton X-100, and 5% goat serum (Gibco 16210-064)) for 1h at room temperature. Slides were then incubated with primary antibody at 4°C overnight in blocking buffer diluted 2-fold with PBS (Supplemental Table 4). Sections were washed 3x for 5m in PBS, then incubated at room temperature for 1h with Alexa Fluor goat anti-mouse 488 (Life Technologies AB150113), Alexa Fluor donkey anti-goat 647 (Life Technologies AB150131), and Alexa Fluor goat anti-rabbit 568 (Life Technologies AB175470) secondary antibodies diluted 1:500 in blocking solution. Sections were then washed 3x for 5m, counterstained, and mounted with Vectashield Hardset with DAPI (Vector Labs H-1500). Images were acquired using Olympus Whole Slide Scanner (VS120) with a 40x objective.

3.5.4 Differentiation of iPSC-derived astrocytes

iPSC-derived astrocyte progenitors were derived, cultured, and expanded as described previously[142]. The glial progenitors were pre-differentiated in Neurobasal Medium containing 1% B27 Supplement, 1% NEAA, and 1% PS for 10d on 6-well plates before cryopreservation. Before the experiment, batches of pre-differentiated astrocytes were defrosted and plated at 50k per well in 8-well Ibidi imaging chambers (Ibidi 80841), and cultured for an additional 6 days in Neurobasal Medium containing 1% B27, 1% NEAA, 1% PS, and 20 ng/mL ciliary neurotrophic factor (CNTF, Thermo Fisher PHC7015). Astrocytes were then fixed in 4% PFA for 15m and immunostained as described above (Supplemental Table 4) using Alexa Fluor goat anti-rabbit 647 (Thermo Fisher A-21245) and Alexa Fluor goat anti-mouse IgG1 (Thermo Fisher A-21121) secondary antibodies

diluted 1:1000 in 3% goat serum in PBS. Imaging was performed on a Leica DMi8 with CooLED light source at 63x and analyzed with ImageJ.

3.6 Ethics statement

All vertebrate animal work was approved by the Committee on the Use and Care of Animals (UCUCA) at the University of Michigan and in accordance with the United Kingdom Animals Act (1986). All experiments were performed in accordance with UCUCA guidelines. Rats (*Rattus norvegicus*) used for primary neuron collection were housed singly in chambers equipped with environmental enrichment. All studies were designed to minimize animal use. Rats were cared for by the Unit for Laboratory Animal Medicine at the University of Michigan; all individuals were trained and approved in the care and long-term maintenance of rodent colonies, in accordance with the NIH-supported Guide for the Care and Use of Laboratory Animals. All personnel handling the rats and administering euthanasia were properly trained in accordance with the UM Policy for Education and Training of Animal Care and Use Personnel. Euthanasia was fully consistent with the recommendations of the Guidelines on Euthanasia of the American Veterinary Medical Association.

3.7 Acknowledgements

This work was supported by National Institutes of Health (NIH), National Institute for Neurological Disorders and Stroke (NINDS) R01-NS097542, National Institute for Aging (NIA) P30 AG053760 (SJB), the University of Michigan Protein Folding Disease Initiative, and Ann Arbor Active Against ALS. We thank Dr. M. Uhler for advice, protocols and reagents; Drs. M. Ward, Z. Xu, and Y. Ayala for reagents; and Dr. J. Parent, Mr. M. Malik, and Mr. M. McMillian

for their suggestions. We also thank Mr. M. Perkins from the Michigan Brain Bank (5P30 AG053760, University of Michigan Alzheimer's Disease Core Center), the Michigan ALS Biorepository, the University of Michigan DNA Sequencing Core, and the University of Michigan Department of Pharmacology for access to their confocal microscopy core. Finally, we thank those that donated the fibroblast and tissue samples that made these studies possible.

3.8 Author Contributions

K.W. was responsible for conceptualization, methodology, investigation, formal analysis, writing and visualization. S.J.B. contributed to conceptualization, methodology, formal analysis, writing, visualization, supervision, project administration, and funding acquisition. E.M.T. contributed to conceptualization and methodology, and R.M. was responsible for software. M.A.W. and N.B.G. contributed to data curation and formal analysis. J.S. contributed to supervision and project administration. J.P.M., Z.L., C.M.G., and A.S. contributed to investigation.

References

1. Bruijn, L. I., Miller, T. M. & Cleveland, D. W. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu. Rev. Neurosci.* **27**, 723–749 (2004).
2. Neumann, M. Molecular neuropathology of TDP-43 proteinopathies. *Int. J. Mol. Sci.* **10**, 232–246 (2009).
3. Murray, M. E. *et al.* Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. *Acta Neuropathol.* **122**, 673–690 (2011).
4. Seilhean, D. *et al.* Accumulation of TDP-43 and α-actin in an amyotrophic lateral sclerosis patient with the K17I ANG mutation. *Acta Neuropathol.* **118**, 561–573 (2009).
5. Van Mossevelde, S. *et al.* Clinical features of TBK1 carriers compared with C9orf72, GRN and non-mutation carriers in a Belgian cohort. *Brain* **139**, 452–467 (2016).
6. Smith, B. N. *et al.* Novel mutations support a role for Profilin 1 in the pathogenesis of ALS. *Neurobiol. Aging* **36**, 1602.e17–27 (2015).
7. Deng, H.-X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).
8. Johnson, J. O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).
9. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).
10. Kraemer, B. C. *et al.* Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis. *Acta Neuropathol.* **119**, 409–419 (2010).
11. Wu, L.-S., Cheng, W.-C. & Shen, C.-K. J. Targeted depletion of TDP-43 expression in the spinal cord motor neurons leads to the development of amyotrophic lateral sclerosis-like phenotypes in mice. *J. Biol. Chem.* **287**, 27335–27344 (2012).
12. Iguchi, Y. *et al.* Loss of TDP-43 causes age-dependent progressive motor neuron degeneration. *Brain* **136**, 1371–1382 (2013).
13. Sephton, C. F. *et al.* TDP-43 is a developmentally regulated protein essential for early embryonic development. *J. Biol. Chem.* **285**, 6826–6834 (2010).
14. Barmada, S. J. *et al.* Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
15. Archbold, H. C. *et al.* TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci. Rep.* **8**, 4606 (2018).
16. Swarup, V. *et al.* Pathological hallmarks of amyotrophic lateral sclerosis/frontotemporal lobar degeneration in transgenic mice produced with TDP-43 genomic fragments. *Brain* **134**, 2610–2626 (2011).
17. Wils, H. *et al.* TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3858–3863 (2010).
18. Dayton, R. D. *et al.* Selective forelimb impairment in rats expressing a pathological TDP-43 25 kDa C-terminal fragment to mimic amyotrophic lateral sclerosis. *Mol. Ther.* **21**, 1324–1334 (2013).
19. Tatom, J. B. *et al.* Mimicking aspects of frontotemporal lobar degeneration and Lou Gehrig's disease in rats via TDP-43 overexpression. *Mol. Ther.* **17**, 607–613 (2009).

20. Voigt, A. *et al.* TDP-43-mediated neuron loss in vivo requires RNA-binding activity. *PLoS One* **5**, e12247 (2010).
21. Li, Y. *et al.* A Drosophila model for TDP-43 proteinopathy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3169–3174 (2010).
22. Kabashi, E. *et al.* Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum. Mol. Genet.* **19**, 671–683 (2010).
23. Schmid, B. *et al.* Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4986–4991 (2013).
24. Uchida, A. *et al.* Non-human primate model of amyotrophic lateral sclerosis with cytoplasmic mislocalization of TDP-43. *Brain* **135**, 833–846 (2012).
25. Jackson, K. L. *et al.* Initial gene vector dosing for studying symptomatology of amyotrophic lateral sclerosis in non-human primates. *J. Med. Primatol.* **44**, 66–75 (2015).
26. Vucic, S., Nicholson, G. A. & Kiernan, M. C. Cortical hyperexcitability may precede the onset of familial amyotrophic lateral sclerosis. *Brain* **131**, 1540–1550 (2008).
27. Kanai, K., Shibuya, K. & Kuwabara, S. Motor axonal excitability properties are strong predictors for survival in amyotrophic lateral sclerosis. *Rinsho Shinkeigaku* **51**, 1118–1119 (2011).
28. Kanai, K. *et al.* Altered axonal excitability properties in amyotrophic lateral sclerosis: impaired potassium channel function related to disease stage. *Brain* **129**, 953–962 (2006).
29. Enterzari-Taher, M., Eisen, A., Stewart, H. & Nakajima, M. Abnormalities of cortical inhibitory neurons in amyotrophic lateral sclerosis. *Muscle Nerve* **20**, 65–71 (1997).
30. Zanette, G. *et al.* Different mechanisms contribute to motor cortex hyperexcitability in amyotrophic lateral sclerosis. *Clin. Neurophysiol.* **113**, 1688–1697 (2002).
31. Geevasinga, N. *et al.* Cortical Function in Asymptomatic Carriers and Patients With C9orf72 Amyotrophic Lateral Sclerosis. *JAMA Neurol.* **72**, 1268–1274 (2015).
32. Vucic, S., Cheah, B. C. & Kiernan, M. C. Defining the mechanisms that underlie cortical hyperexcitability in amyotrophic lateral sclerosis. *Exp. Neurol.* **220**, 177–182 (2009).
33. Karandreas, N. *et al.* Impaired interhemispheric inhibition in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **8**, 112–118 (2007).
34. Tamura, N. *et al.* Increased nodal persistent Na currents in human neuropathy and motor neuron disease estimated by latent addition. *Clinical Neurophysiology* **117**, 2451–2458 (2006).
35. Shibuya, K. *et al.* Markedly reduced axonal potassium channel expression in human sporadic amyotrophic lateral sclerosis: An immunohistochemical study. *Experimental Neurology* **232**, 149–153 (2011).
36. Geevasinga, N. *et al.* Riluzole exerts transient modulating effects on cortical and axonal hyperexcitability in ALS. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **17**, 580–588 (2016).
37. Kuo, J. J. *et al.* Hyperexcitability of cultured spinal motoneurons from presymptomatic ALS mice. *J. Neurophysiol.* **91**, 571–575 (2004).
38. Pieri, M. *et al.* Altered excitability of motor neurons in a transgenic mouse model of familial amyotrophic lateral sclerosis. *Neurosci. Lett.* **351**, 153–156 (2003).
39. Van Zundert, B. *et al.* Neonatal neuronal circuitry shows hyperexcitable disturbance in a mouse model of the adult-onset neurodegenerative disease amyotrophic lateral sclerosis. *Journal of Neuroscience* **28**, 10864–10874 (2008).

40. Saba, L. *et al.* Altered Functionality, Morphology, and Vesicular Glutamate Transporter Expression of Cortical Motor Neurons from a Presymptomatic Mouse Model of Amyotrophic Lateral Sclerosis. *Cerebral Cortex* **26**, 1512–1528 (2016).
41. Pieri, M., Carunchio, I., Curcio, L., Mercuri, N. B. & Zona, C. Increased persistent sodium current determines cortical hyperexcitability in a genetic model of amyotrophic lateral sclerosis. *Exp. Neurol.* **215**, 368–379 (2009).
42. Wainger, B. J. *et al.* Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep.* **7**, 1–11 (2014).
43. Devlin, A.-C. *et al.* Human iPSC-derived motoneurons harbouring TARDBP or C9ORF72 ALS mutations are dysfunctional despite maintaining viability. *Nat. Commun.* **6**, 5999 (2015).
44. Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci. Transl. Med.* **5**, 208ra149 (2013).
45. Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78**, 785–798 (2013).
46. Lam, R. S., Töpfer, F. M., Wood, P. G., Busskamp, V. & Bamberg, E. Functional Maturation of Human Stem Cell-Derived Neurons in Long-Term Cultures. *PLoS One* **12**, e0169506 (2017).
47. Busskamp, V. *et al.* Rapid neurogenesis through transcriptional activation in human stem cells. *Mol. Syst. Biol.* **10**, 760 (2014).
48. Gupta, S. *et al.* Fibroblast growth factor 2 regulates activity and gene expression of human post-mitotic excitatory neurons. *J. Neurochem.* **145**, 188–203 (2018).
49. Wang, I.-F., Wu, L.-S., Chang, H.-Y. & Shen, C.-K. J. TDP-43, the signature protein of FTLD-U, is a neuronal activity-responsive factor. *J. Neurochem.* **105**, 797–806 (2008).
50. Schwenk, B. M. *et al.* TDP-43 loss of function inhibits endosomal trafficking and alters trophic signaling in neurons. *The EMBO Journal* **35**, 2350–2370 (2016).
51. Gopal, P. P., Nirschl, J. J., Klinman, E. & Holzbaur, E. L. F. Amyotrophic lateral sclerosis-linked mutations increase the viscosity of liquid-like TDP-43 RNP granules in neurons. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E2466–E2475 (2017).
52. Zhang, Y.-J. *et al.* The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112–3122 (2013).
53. Suzuki, H., Shibagaki, Y., Hattori, S. & Matsuoka, M. Nuclear TDP-43 causes neuronal toxicity by escaping from the inhibitory regulation by hnRNPs. *Hum. Mol. Genet.* **24**, 1513–1527 (2015).
54. Baker, M. Reproducibility crisis: Blame it on the antibodies. *Nature* **521**, 274–276 (2015).
55. Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
56. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
57. Eom, T. *et al.* NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife* **2**, e00178 (2013).
58. Kolisnyk, B. *et al.* Cholinergic Regulation of hnRNPA2/B1 Translation by M1 Muscarinic Receptors. *J. Neurosci.* **36**, 6287–6296 (2016).
59. Flores, B. N. *et al.* An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep.* **27**, 1133–1150.e8 (2019).
60. Wang, H.-Y., Wang, I.-F., Bose, J. & Shen, C.-K. J. Structural diversity and functional

implications of the eukaryotic TDP gene family. *Genomics* **83**, 130–139 (2004).
61. D'Alton, S., Altshuler, M. & Lewis, J. Studies of alternative isoforms provide insight into TDP-43 autoregulation and pathogenesis. *RNA* **21**, 1419–1432 (2015).
62. Avendaño-Vázquez, S. E. *et al.* Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev.* **26**, 1679–1684 (2012).
63. Bembich, S. *et al.* Predominance of spliceosomal complex formation over polyadenylation site selection in TDP-43 autoregulation. *Nucleic Acids Res.* **42**, 3362–3371 (2014).
64. Seyfried, N. T. *et al.* Multiplex SILAC analysis of a cellular TDP-43 proteinopathy model reveals protein inclusions associated with SUMOylation and diverse polyubiquitin chains. *Mol. Cell. Proteomics* **9**, 705–718 (2010).
65. Dammer, E. B. *et al.* Coaggregation of RNA-binding proteins in a model of TDP-43 proteinopathy with selective RGG motif methylation and a role for RRM1 ubiquitination. *PLoS One* **7**, e38658 (2012).
66. la Cour, T. *et al.* Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng. Des. Sel.* **17**, 527–536 (2004).
67. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic control of nuclear protein export. *Nature Communications* **7**, (2016).
68. Ayala, Y. M. *et al.* Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778–3785 (2008).
69. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
70. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *Elife* **7**, (2018).
71. Weskamp, K., Safren, N., Miguez, R. & Barmada, S. Monitoring Neuronal Survival via Longitudinal Fluorescence Microscopy. *JoVE (Journal of Visualized Experiments)* e59036 (2019).
72. Christensen, E. Multivariate survival analysis using Cox's regression model. *Hepatology* **7**, 1346–1358 (1987).
73. Afroz, T. *et al.* Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature Communications* **8**, (2017).
74. Kuo, P.-H., Doudeva, L. G., Wang, Y.-T., Shen, C.-K. J. & Yuan, H. S. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* **37**, 1799–1808 (2009).
75. Johnson, B. S. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* **284**, 20329–20339 (2009).
76. Zhang, Y.-J. *et al.* Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7607–7612 (2009).
77. Bozzo, F. *et al.* Structural insights into the multi-determinant aggregation of TDP-43 in motor neuron-like cells. *Neurobiol. Dis.* **94**, 63–72 (2016).
78. Budini, M., Romano, V., Quadri, Z., Buratti, E. & Baralle, F. E. TDP-43 loss of cellular function through aggregation requires additional structural determinants beyond its C-terminal Q/N prion-like domain. *Hum. Mol. Genet.* **24**, 9–20 (2015).
79. Chang, C.-K. *et al.* The N-terminus of TDP-43 promotes its oligomerization and enhances DNA binding affinity. *Biochem. Biophys. Res. Commun.* **425**, 219–224 (2012).

80. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R. & Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810 (2004).
81. Ayala, Y. M., Pagani, F. & Baralle, F. E. TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS Lett.* **580**, 1339–1344 (2006).
82. Buratti, E. *et al.* Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J.* **20**, 1774–1784 (2001).
83. 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).
84. Ayala, Y. M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).
85. White, M. A. *et al.* TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat. Neurosci.* **21**, 552–563 (2018).
86. Krach, F. *et al.* Transcriptome–pathology correlation identifies interplay between TDP-43 and the expression of its kinase CK1E in sporadic ALS. *Acta Neuropathologica* **136**, 405–423 (2018).
87. D'Erchia, A. M. *et al.* Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS. *Sci. Rep.* **7**, 10046 (2017).
88. Prudencio, M. *et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175–1182 (2015).
89. Li, H.-R., Chiang, W.-C., Chou, P.-C., Wang, W.-J. & Huang, J.-R. TAR DNA-binding protein 43 (TDP-43) liquid–liquid phase separation is mediated by just a few aromatic residues. *J. Biol. Chem.* **293**, 6090–6098 (2018).
90. McGurk, L. *et al.* Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization. *Molecular Cell* **71**, 703–717.e9 (2018).
91. Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537–1549 (2016).
92. Mann, J. R. *et al.* RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron* (2019). doi:10.1016/j.neuron.2019.01.048
93. Agrawal, S. *et al.* RNA recognition motifs of disease-linked RNA-binding proteins contribute to amyloid formation. *Sci. Rep.* **9**, 6171 (2019).
94. Franzmann, T. & Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* (2018). doi:10.1074/jbc.TM118.001190
95. Riback, J. A. *et al.* Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **168**, 1028–1040.e19 (2017).
96. Kroschwald, S. *et al.* Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. *Cell Rep.* **23**, 3327–3339 (2018).
97. Franzmann, T. M. *et al.* Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, (2018).
98. Winton, M. J. *et al.* Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *J. Biol. Chem.* **283**, 13302–13309 (2008).
99. Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I. W. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051–1060 (1997).

100. Pinarbasi, E. S. *et al.* Active nuclear import and passive nuclear export are the primary determinants of TDP-43 localization. *Sci. Rep.* **8**, 7083 (2018).
101. Kırlı, K. *et al.* A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *Elife* **4**, (2015).
102. Thakar, K., Karaca, S., Port, S. A., Urlaub, H. & Kehlenbach, R. H. Identification of CRM1-dependent Nuclear Export Cargos Using Quantitative Mass Spectrometry. *Mol. Cell. Proteomics* **12**, 664–678 (2013).
103. Ayala, Y. M. *et al.* Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol.* **348**, 575–588 (2005).
104. Johnson, B. S., McCaffery, J. M., Lindquist, S. & Gitler, A. D. A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6439–6444 (2008).
105. Wang, A., Conicella, A. E. & Schmidt, H. B. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J.* (2018).
106. Leigh, P. N. *et al.* Ubiquitin-immunoreactive intraneuronal inclusions in amyotrophic lateral sclerosis. Morphology, distribution, and specificity. *Brain* **114 (Pt 2)**, 775–788 (1991).
107. Ravits, J., Paul, P. & Jorg, C. Focality of upper and lower motor neuron degeneration at the clinical onset of ALS. *Neurology* **68**, 1571–1575 (2007).
108. Udan, M. & Baloh, R. H. Implications of the prion-related Q/N domains in TDP-43 and FUS. *Prion* **5**, 1–5 (2011).
109. Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K. & Nukina, N. A seeding reaction recapitulates intracellular formation of Sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. *J. Biol. Chem.* **286**, 18664–18672 (2011).
110. Nonaka, T., Kametani, F., Arai, T., Akiyama, H. & Hasegawa, M. Truncation and pathogenic mutations facilitate the formation of intracellular aggregates of TDP-43. *Hum. Mol. Genet.* **18**, 3353–3364 (2009).
111. Tsuji, H. *et al.* Molecular analysis and biochemical classification of TDP-43 proteinopathy. *Brain* **135**, 3380–3391 (2012).
112. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
113. Igaz, L. M. *et al.* Expression of TDP-43 C-terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies. *J. Biol. Chem.* **284**, 8516–8524 (2009).
114. Wang, I.-F. *et al.* The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. *Nat. Commun.* **3**, 766 (2012).
115. Wang, Y.-T. *et al.* The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J. Biol. Chem.* **288**, 9049–9057 (2013).
116. Berning, B. A. & Walker, A. K. The Pathobiology of TDP-43 C-Terminal Fragments in ALS and FTLD. *Front. Neurosci.* **13**, 335 (2019).
117. Yamashita, T. *et al.* A role for calpain-dependent cleavage of TDP-43 in amyotrophic lateral sclerosis pathology. *Nat. Commun.* **3**, 1307 (2012).
118. Igaz, L. M. *et al.* Enrichment of C-Terminal Fragments in TAR DNA-Binding Protein-43 Cytoplasmic Inclusions in Brain but not in Spinal Cord of Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *The American Journal of Pathology* **173**, 182–194 (2008).
119. Zacco, E., Martin, S. R., Thorogate, R. & Pastore, A. The RNA-Recognition Motifs of TAR

DNA-Binding Protein 43 May Play a Role in the Aberrant Self-Assembly of the Protein. *Frontiers in Molecular Neuroscience* **11**, (2018).
120. Sasaguri, H. *et al.* The extreme N-terminus of TDP-43 mediates the cytoplasmic aggregation of TDP-43 and associated toxicity in vivo. *Brain Res.* **1647**, 57–64 (2016).
121. Shodai, A. *et al.* Aberrant Assembly of RNA Recognition Motif 1 Links to Pathogenic Conversion of TAR DNA-binding Protein of 43 kDa (TDP-43). *Journal of Biological Chemistry* **288**, 14886–14905 (2013).
122. Neumann, M. *et al.* TDP-43-Positive White Matter Pathology in Frontotemporal Lobar Degeneration With Ubiquitin-Positive Inclusions. *Journal of Neuropathology and Experimental Neurology* **66**, 177–183 (2007).
123. Nishihira, Y. *et al.* Sporadic amyotrophic lateral sclerosis: two pathological patterns shown by analysis of distribution of TDP-43-immunoreactive neuronal and glial cytoplasmic inclusions. *Acta Neuropathologica* **116**, 169–182 (2008).
124. Zhang, H. *et al.* TDP-43-immunoreactive neuronal and glial inclusions in the neostriatum in amyotrophic lateral sclerosis with and without dementia. *Acta Neuropathologica* **115**, 115–122 (2007).
125. Serio, A. *et al.* Astrocyte pathology and the absence of non-cell autonomy in an induced pluripotent stem cell model of TDP-43 proteinopathy. *Proceedings of the National Academy of Sciences* **110**, 4697–4702 (2013).
126. Bruijn, L. I. *et al.* ALS-Linked SOD1 Mutant G85R Mediates Damage to Astrocytes and Promotes Rapidly Progressive Disease with SOD1-Containing Inclusions. *Neuron* **18**, 327–338 (1997).
127. Guo, H. Increased expression of the glial glutamate transporter EAAT2 modulates excitotoxicity and delays the onset but not the outcome of ALS in mice. *Human Molecular Genetics* **12**, 2519–2532 (2003).
128. Howland, D. S. *et al.* Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proceedings of the National Academy of Sciences* **99**, 1604–1609 (2002).
129. Pardo, A. C. *et al.* Loss of the astrocyte glutamate transporter GLT1 modifies disease in SOD1G93A mice. *Experimental Neurology* **201**, 120–130 (2006).
130. Rothstein, J. D., Martin, L. J. & Kuncl, R. W. Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis. *N. Engl. J. Med.* **326**, 1464–1468 (1992).
131. Rothstein, J. D., Van Kammen, M., Levey, A. I., Martin, L. J. & Kuncl, R. W. Selective loss of glial glutamate transporter GLT-1 in amyotrophic lateral sclerosis. *Ann. Neurol.* **38**, 73–84 (1995).
132. Yu, J., Chau, K. F., Vodyanik, M. A., Jiang, J. & Jiang, Y. Efficient feeder-free episomal reprogramming with small molecules. *PLoS One* **6**, e17557 (2011).
133. Tank, E. M. *et al.* Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat. Commun.* **9**, 2845 (2018).
134. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
135. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
136. Barmada, S. J. *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat. Chem. Biol.* **10**, 677–685 (2014).
137. Saudou, F., Finkbeiner, S., Devys, D. & Greenberg, M. E. Huntingtin Acts in the Nucleus to

Induce Apoptosis but Death Does Not Correlate with the Formation of Intranuclear Inclusions. Cell **95**, 55–66 (1998).
138. Duong, H. & Han, M. A multispectral LED array for the reduction of background autofluorescence in brain tissue. J. Neurosci. Methods 220, 46–54 (2013).
139. Prudencio, M. et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat. Neurosci. **18**, 1175–1182 (2015).
140. D'Erchia, A. M. et al. Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS. Sci. Rep. 7, 10046 (2017).
141. Yang, C. et al. Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. Proc. Natl. Acad. Sci. U. S. A. **111**, E1121–9 (2014).
142. Serio, A. et al. Astrocyte pathology and the absence of non-cell autonomy in an induced pluripotent stem cell model of TDP-43 proteinopathy. Proceedings of the National Academy of Sciences **110**, 4697–4702 (2013).

Figures

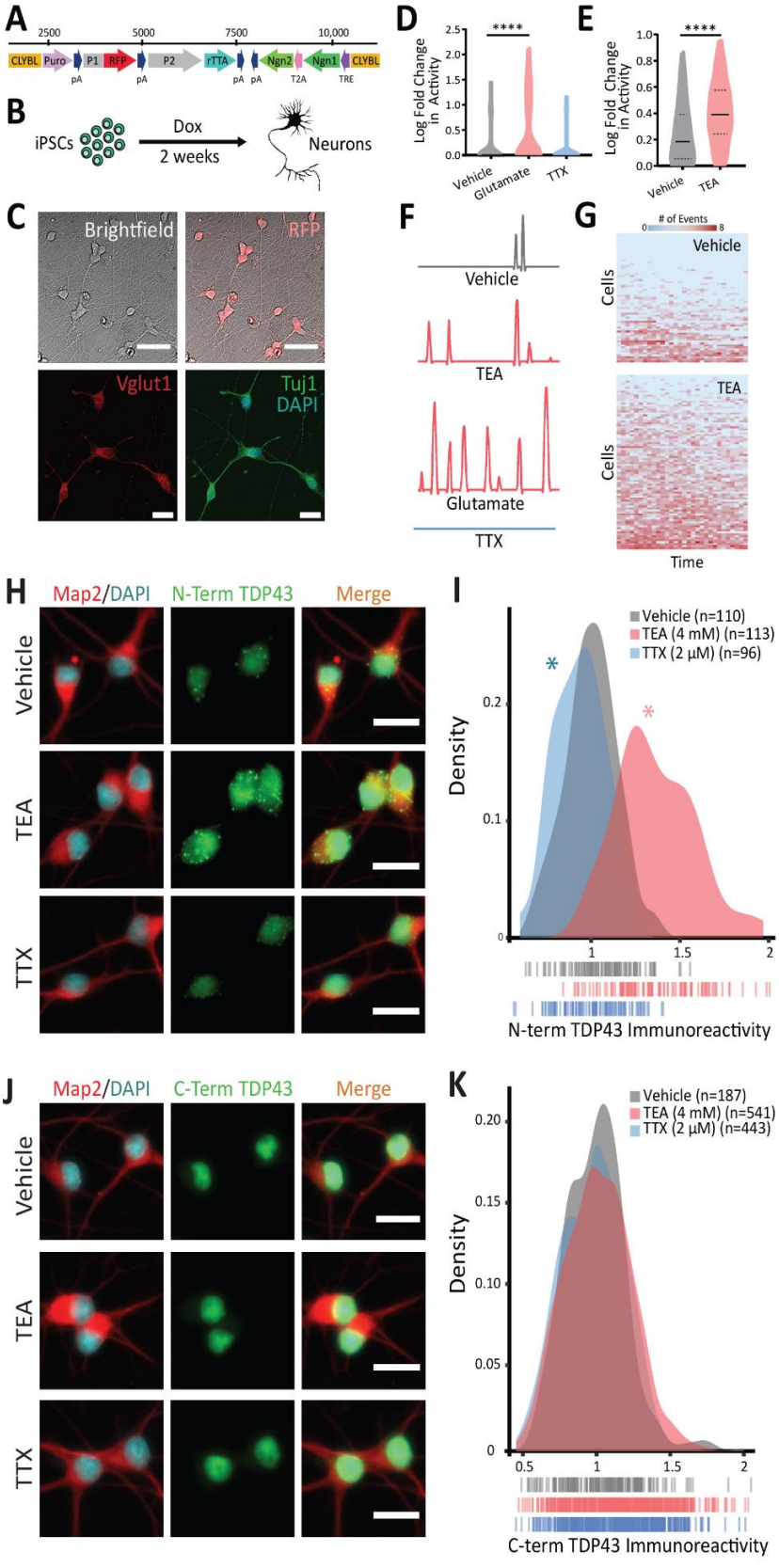

A
2500
5000
7500
10,000
CLYBL
Puro
P1
RFP
P2
rTTA
Ngn2
Ngn1
CLYBL
pA
pA
pA
pA
T2A
TRE
B
iPSCs
Dox
2 weeks
Neurons
C
Brightfield
RFP
Vglut1
Tuj1
DAPI
D
Log Fold Change in Activity
****
Vehicle
Glutamate
TTX
E
Log Fold Change in Activity
****
Vehicle
TEA
F
Vehicle
TEA
Glutamate
TTX
G
# of Events
Vehicle
Cells
TEA
Cells
Time
H
Map2/DAPI
N-Term TDP43
Merge
Vehicle
TEA
TTX
I
Vehicle (n=110)
TEA (4 mM) (n=113)
TTX (2 µM) (n=96)
Density
N-term TDP43 Immunoreactivity
J
Map2/DAPI
C-Term TDP43
Merge
Vehicle
TEA
TTX
K
Vehicle (n=187)
TEA (4 mM) (n=541)
TTX (2 µM) (n=443)
Density
C-term TDP43 Immunoreactivity

Figure 3.1. Hyperexcitability drives TDP43 accumulation in human iNeurons. (A) Schematic of the cassette used to integrate Ngn1 and Ngn2 into the *CLYBL* safe harbor locus under a doxycycline-inducible (Tet-on) promoter. CLYBL, targeting sequence; Puro, puromycin resistance gene; pA, poly-A tail; P1,P2, promoters; RFP, mCherry; rtTA, reverse tetracycline-controlled transactivator; Ngn1,2, Neurogenin1 and 2; T2A, self-cleaving peptide; TRE, tetracycline response element. (B) Timeline depicting the differentiation of iPSCs into forebrain-like neurons within 2w of doxycycline addition. (C) The resultant neurons are RFP-positive and express the neuronal markers Vglut1 and Tuj1. (D) Spontaneous neuronal activity visualized by the Ca^{2+} reporter gCaMP6f at 2w. Activity was pharmacologically modulated with bath application of glutamate or TTX. Vehicle n=257, Glutamate n=327, TTX n=403, stratified among 3 replicates; ****$p<0.0001$, one-way ANOVA with Dunnett's post-test. (E) Treatment with TEA significantly increased neuronal activity. Vehicle n= 312, TEA n=369, stratified among 3 replicates, **** $p<0.0001$, two-tailed t-test. (F) Example traces depicting changes in gCaMP6f fluorescence for each condition. (G) Heat maps depicting global changes in activity. Each row represents one neuron, and each column represents a 20s observation window. Thirty intervals were collected over a 12h period. Box color indicates the relative firing rate of each cell at each timepoint ranging from low (blue) to high (red). (H) N-terminal TDP43 immunoreactivity was increased in TEA-treated iNeurons and decreased in TTX-treated iNeurons (TTX), indicating a bidirectional relationship between activity and TDP43 abundance. (I) Density plot depicting the change in TDP43 immunoreactivity between conditions. Vehicle n=110, TEA n=113, TTX n=96, 2 replicates, dashes indicate single neurons, *$p<0.01$, Kolmogorov-Smirnov test. (J) No such relationship was identified when TDP43 abundance is detected using an antibody directed against the C-terminus. (K) Density plot depicting the change in C-terminal TDP43 immunoreactivity between conditions. Vehicle n=187, TEA n=541, TTX n=443, 2 replicates, dashes indicate single neurons, not significant by the Kolmogorov-Smirnov test. Scale bars in (C), 50 µm top, 20 µm bottom. Scale bars in (H), (J), 20 µm.

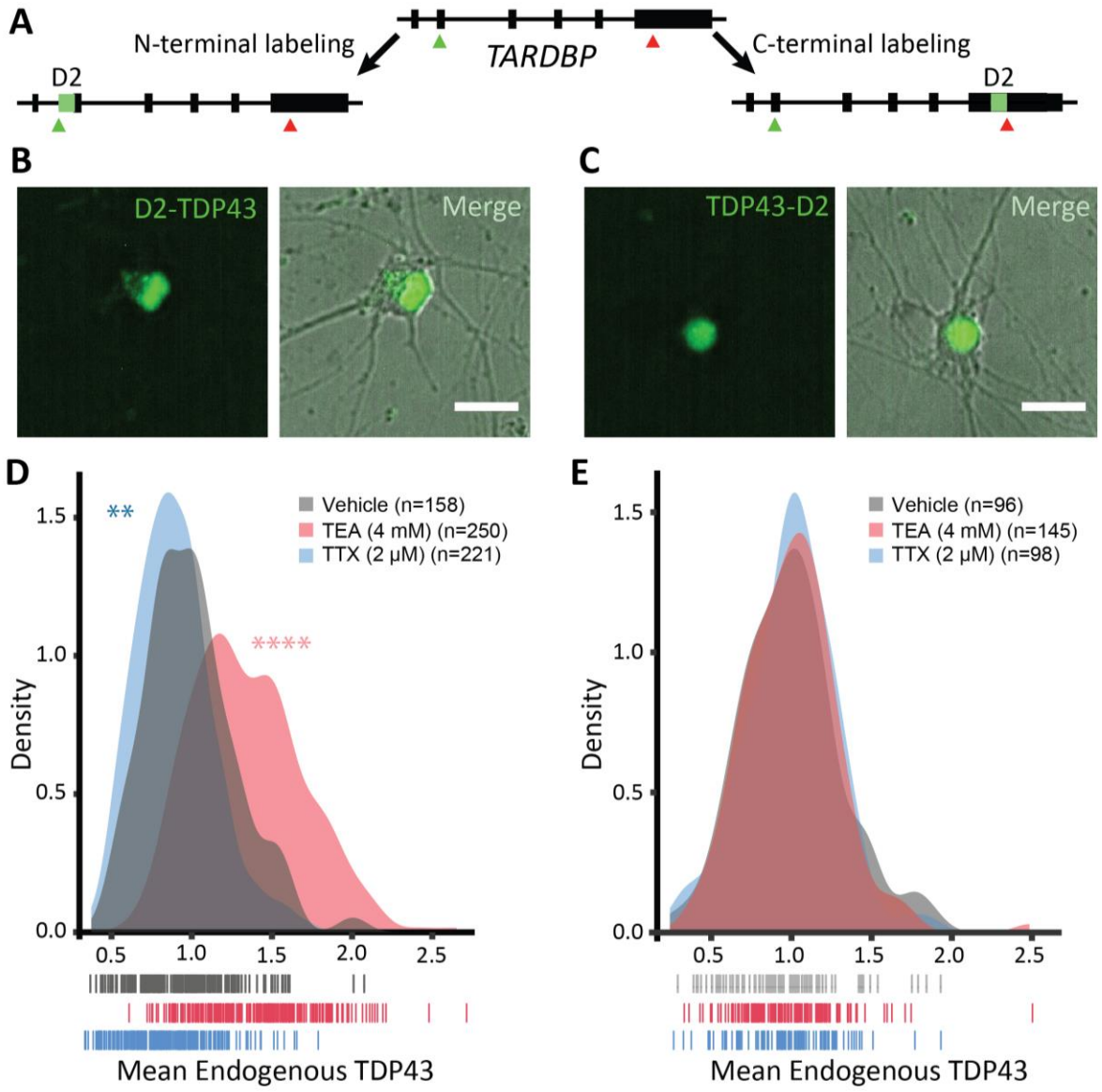


Figure 3.2. TDP43 species harboring the N- but not the C-terminus are regulated by neuronal activity. (A) Strategy for labeling native TDP43 in human iPSC-derived neurons using CRISPR/Cas9. Dendra2 (D2, green) was inserted 3' to the *TARDBP* start codon (green arrow) or 5' to the conventional stop codon (red arrow), enabling fluorescent labeling of the TDP43 N- or C-terminus, respectively. In iNeurons, N-terminally tagged TDP43 (B, D2-TDP43) appeared both nuclear and cytoplasmic in distribution, while C-terminally tagged TDP43 (C, TDP43-D2) was primarily nuclear. (D) Density plot depicting the fluorescence intensity of D2-TDP43 upon application of vehicle (n=158), TEA (n=250), or TTX (n=221). (E) Density plot depicting the fluorescence intensity of TDP43-D2 with addition of vehicle (n=96), TEA (n=145), or TTX (n=98). In (D) and (E), dashes indicate individual neurons from 2 replicates, **p<0.01, ****p<0.0001, Kolmogorov-Smirnov test. Scale bars in (B) and (C), 20 µm.

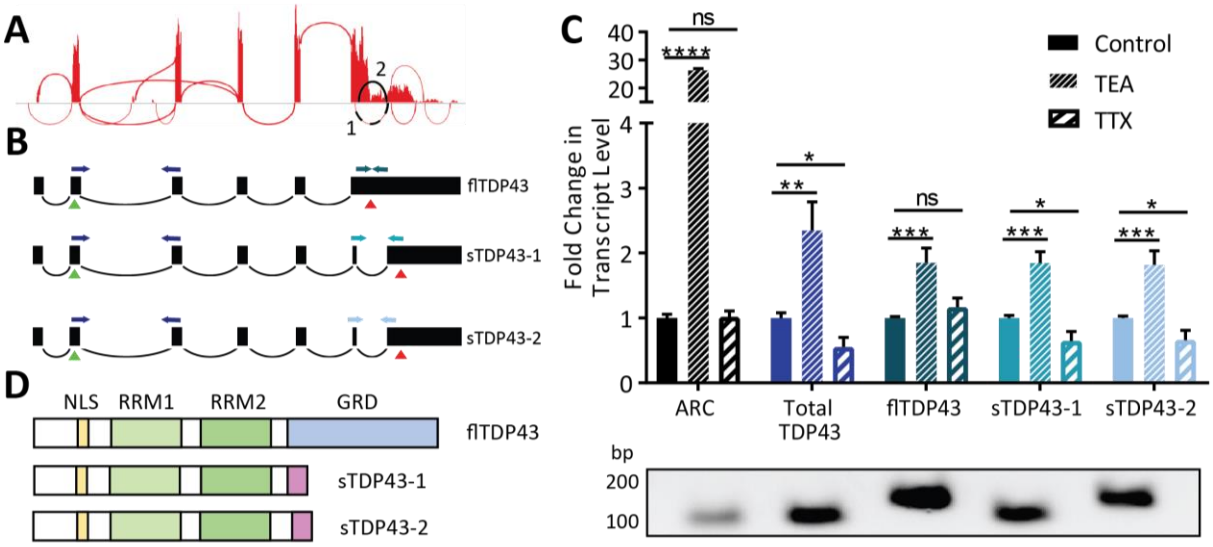


Figure 3.3. Hyperactivity drives alternative splicing of *TARDBP*. (A) Sashimi plot depicting splicing events for the *TARDBP* gene, assembled from HEK293T cell RNA-seq data[59]. Splicing events predicted to skip the majority of exon 6—encoding the TDP43 C-terminus—are highlighted in black. (B) Schematic of transcripts predicted to result in full-length (fl) TDP43 and C-terminally shortened (s) TDP43. Green triangles indicate start codons, red triangles indicate stop codons, and PCR primers are color-coded. (C) qRT-PCR of human iNeurons treated with TEA or TTX, showing activity dependent upregulation of total and sTDP43 or downregulation of sTDP43, respectively. ARC (activity related cytoskeleton associated protein) acts as a positive control for activity-dependent gene regulation. PCR products corresponding to each primer pair are shown below. Data were combined from 3 replicates, **p<0.01, ***p<0.001, ****p<0.0001, two-tailed t-test. (D) Schematic comparing flTDP43 and sTDP43 proteins. Novel sTDP43 C-terminus is shown in purple; NLS, nuclear localization signal; RRM, RNA-recognition motif; GRD, glycine rich domain.

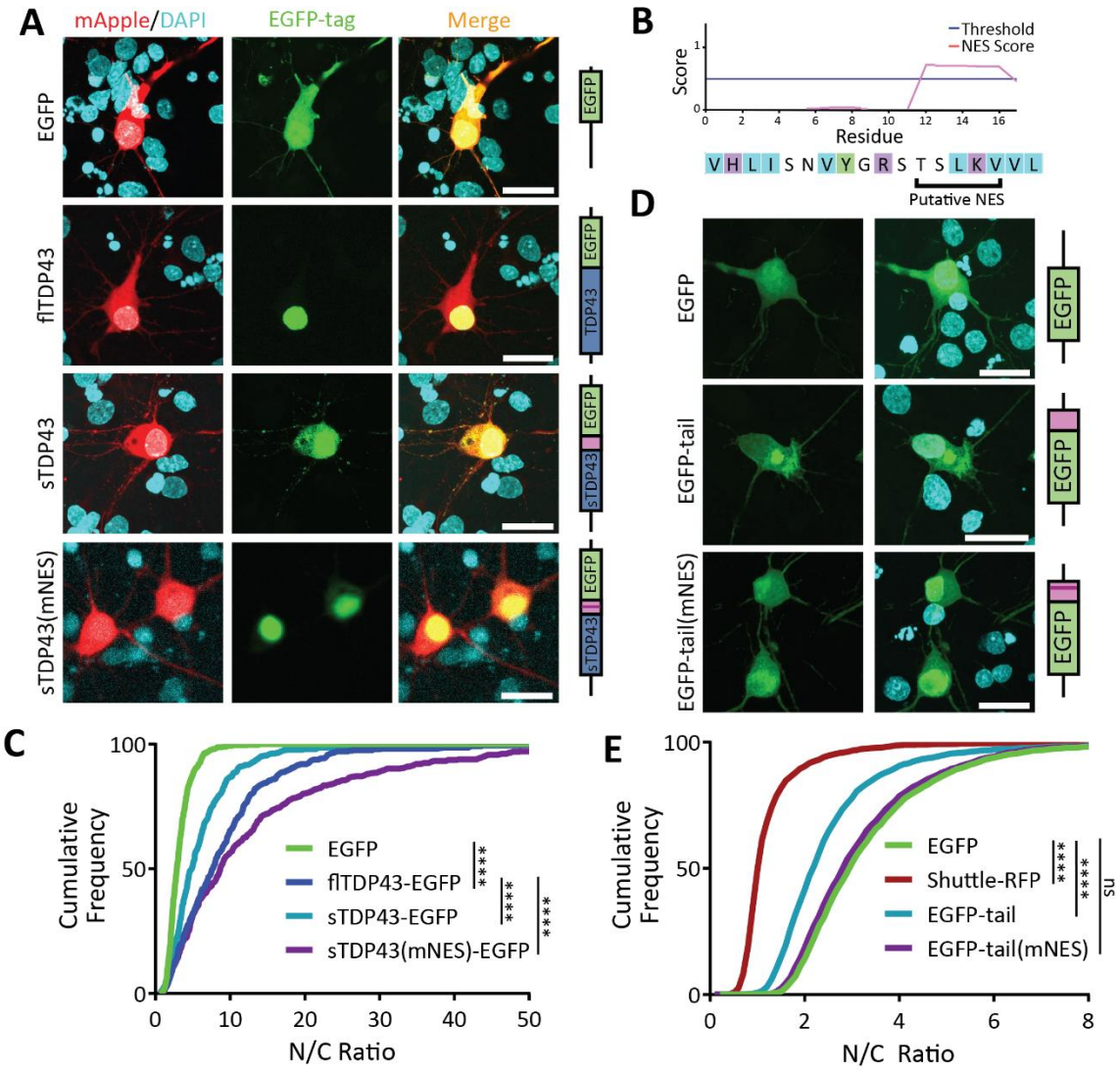


Figure 3.4. sTDP43 accumulates within the cytoplasm due to a putative NES. (A) Rat primary mixed cortical neurons were transfected with mApple and EGFP-tagged TDP43 isoforms, then imaged by fluorescence microscopy. (B) Amino acid sequence of the sTDP43 tail includes a putative NES identified through predictive software NetNES 1.1. Light blue, polar; purple, positively charged; green, hydrophobic residues. (C) sTDP43-EGFP was significantly more cytoplasmic in distribution compared to flTDP43-EGFP, while mutation of the putative NES (mNES) restores nuclear localization. EGFP n=481, flTDP43-EGFP n=385, sTDP43-EGFP n=456, sTDP43(mNES)-EGFP n=490, stratified among 3 replicates, ****p<0.0001, one-way ANOVA with Dunnett's post-test. (D) Rat primary mixed cortical neurons were transfected with EGFP or EGFP fused to either the novel C-terminal tail of sTDP43 or a tail harboring a mutated NES (mNES). (E) The C-terminal sTDP43 tail is sufficient to significantly mislocalize EGFP to the cytoplasm, and mislocalization depends on the NES. Shuttle-RFP, a construct with a strong NES, serves as a positive control for a cytoplasmic protein. EGFP n=2490, Shuttle-RFP n=2073, EGFP-tail n=1956, EGFP-tail(mNES) n=2482, stratified among 3 replicates, ****p<0.0001, one-way ANOVA with Dunnett's post-test. Scale bars in (A) and (D), 20 µm.

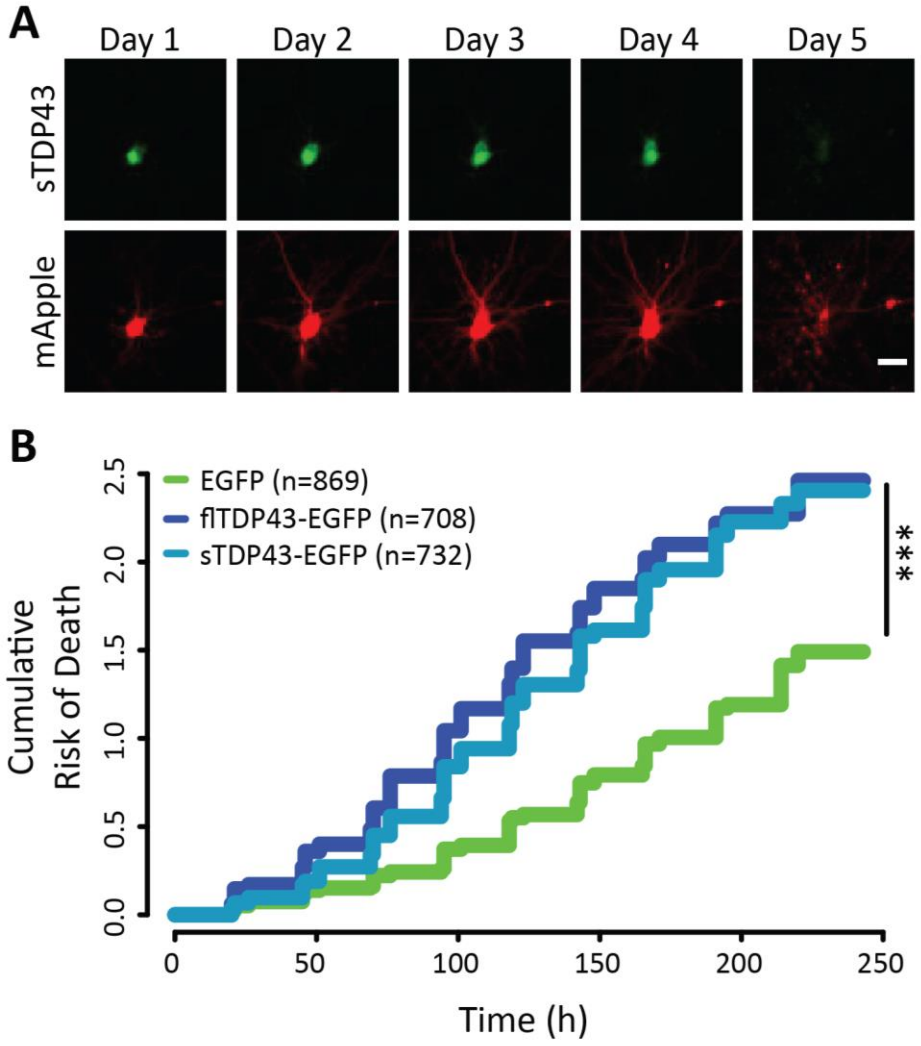


Figure 3.5. sTDP43 overexpression is neurotoxic. (A) Example of a single neuron expressing mApple and sTDP43-EGFP, tracked by longitudinal fluorescence microscopy. Fragmentation of the cell body and loss of fluorescence on Day 5 indicates cell death. (B) The risk of death was significantly greater in neurons overexpressing sTDP43-EGFP and flTDP43-EGFP, in comparison to those expressing EGFP alone. EGFP n= 869, flTDP43-EGFP n=708, sTDP43-EGFP n=732, stratified among 3 replicates, ***$p<2x10^{-16}$, Cox proportional hazards analysis. Scale bar in (A), 20 µm.

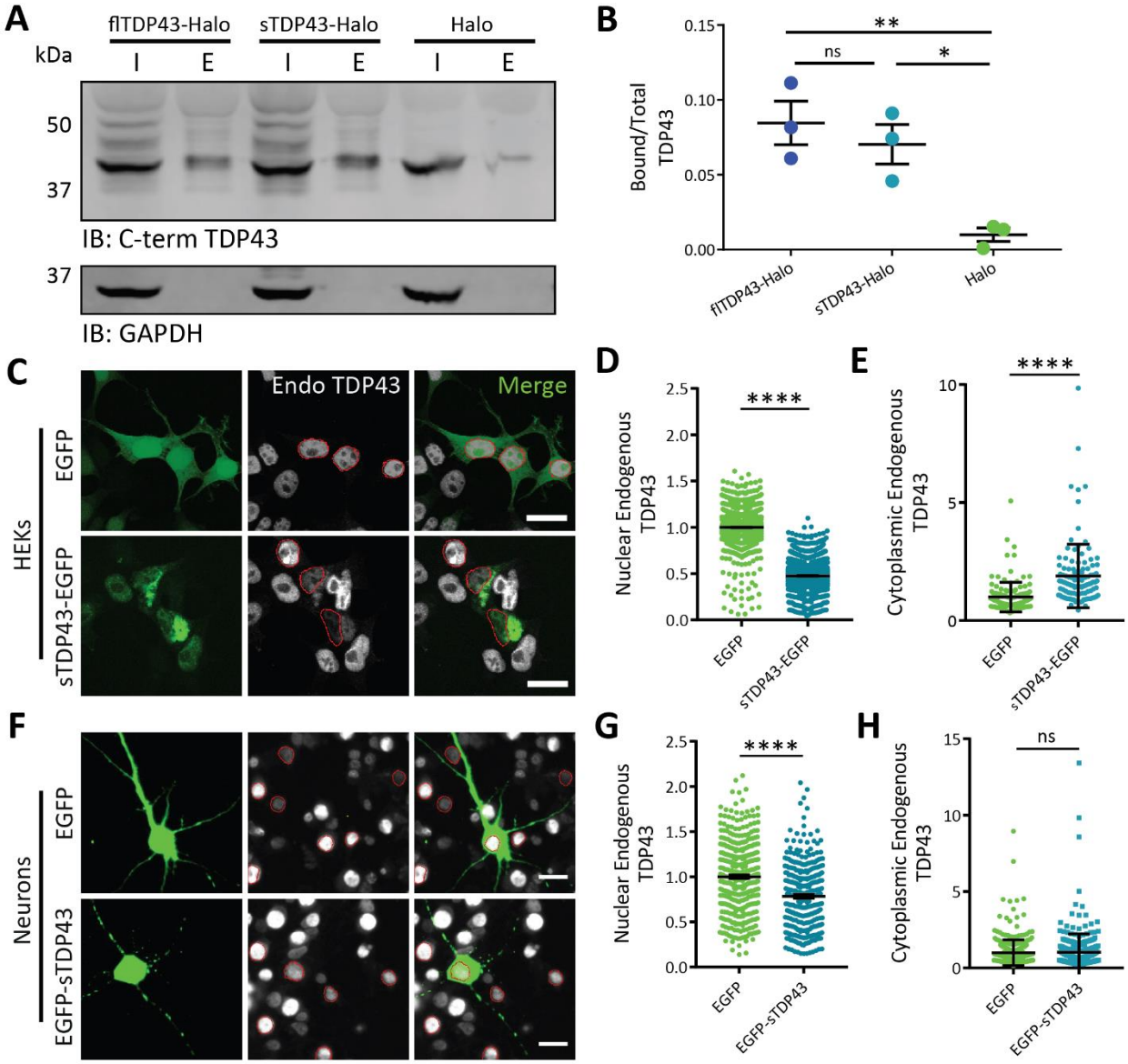

A
flTDP43-Halo
sTDP43-Halo
Halo
kDa
I
E
I
E
I
E
50
37
IB: C-term TDP43
37
IB: GAPDH
B
Bound/Total TDP43
0.15
0.10
0.05
0.00
**
ns
*
flTDP43-Halo
sTDP43-Halo
Halo
C
Endo TDP43
Merge
HEKs
EGFP
sTDP43-EGFP
D
Nuclear Endogenous TDP43
2.5
2.0
1.5
1.0
0.5
0.0
****
EGFP
sTDP43-EGFP
E
Cytoplasmic Endogenous TDP43
10
5
0
****
EGFP
sTDP43-EGFP
F
Neurons
EGFP
EGFP-sTDP43
G
Nuclear Endogenous TDP43
2.5
2.0
1.5
1.0
0.5
0.0
****
EGFP
EGFP-sTDP43
H
Cytoplasmic Endogenous TDP43
15
10
5
0
ns
EGFP
EGFP-sTDP43

Figure 3.6. sTDP43 overexpression leads to the cytoplasmic deposition and nuclear clearance of endogenous TDP43. (A) HaloTagged flTDP43 or sTDP43 were expressed in HEK293T cells and immunoprecipitated with HaloLink. Bound TDP43 was immunoblotted with a C-terminal TDP43 antibody. GAPDH served as a loading control. Input, (I); eluate, (E). (B) Quantification of data shown in (A), demonstrating the fraction of total TDP43 bound to flTDP43-Halo, sTDP43-Halo, or Halo alone. Data was combined from 3 replicates, *p<0.05, **p<0.01, one-way ANOVA with Dunnett's post-test. (C) HEK293T cells were transfected with EGFP or EGFP-tagged sTDP43, then immunostained using an antibody that recognizes the C-terminus of endogenous (Endo) TDP43. Red, nuclear regions of interest (ROIs) determined by DAPI staining. (D) Nuclear, endogenous TDP43 is reduced by sTDP43 overexpression in HEK293T cells. EGFP n=1537, sTDP43-EGFP n=1997, 3 replicates, ****p<0.0001, two-tailed t-test. (E) Cytoplasmic endogenous TDP43 is elevated by sTDP43 overexpression in HEK293T cells. EGFP n=129, sTDP43-EGFP n=113, 3 replicates, ****p<0.0001, two-tailed t-test. (F) Primary mixed rat cortical neurons were transfected with EGFP or EGFP-tagged sTDP43, then immunostained using a C-terminal TDP43 antibody. Red, nuclear ROIs determined by DAPI staining. (G) sTDP43 overexpression resulted in a significant drop in nuclear, endogenous TDP43 in primary neurons (EGFP n=395, EGFP-sTDP43 n=323, 3 replicates, ****p<0.0001, two-tailed t-test), but this was not accompanied by increases in cytoplasmic, endogenous TDP43 (H) (EGFP=394, EGFP-sTDP43=323, 3 replicates, ns by two-tailed t-test). Scale bar in (C), (F) 20 μm.

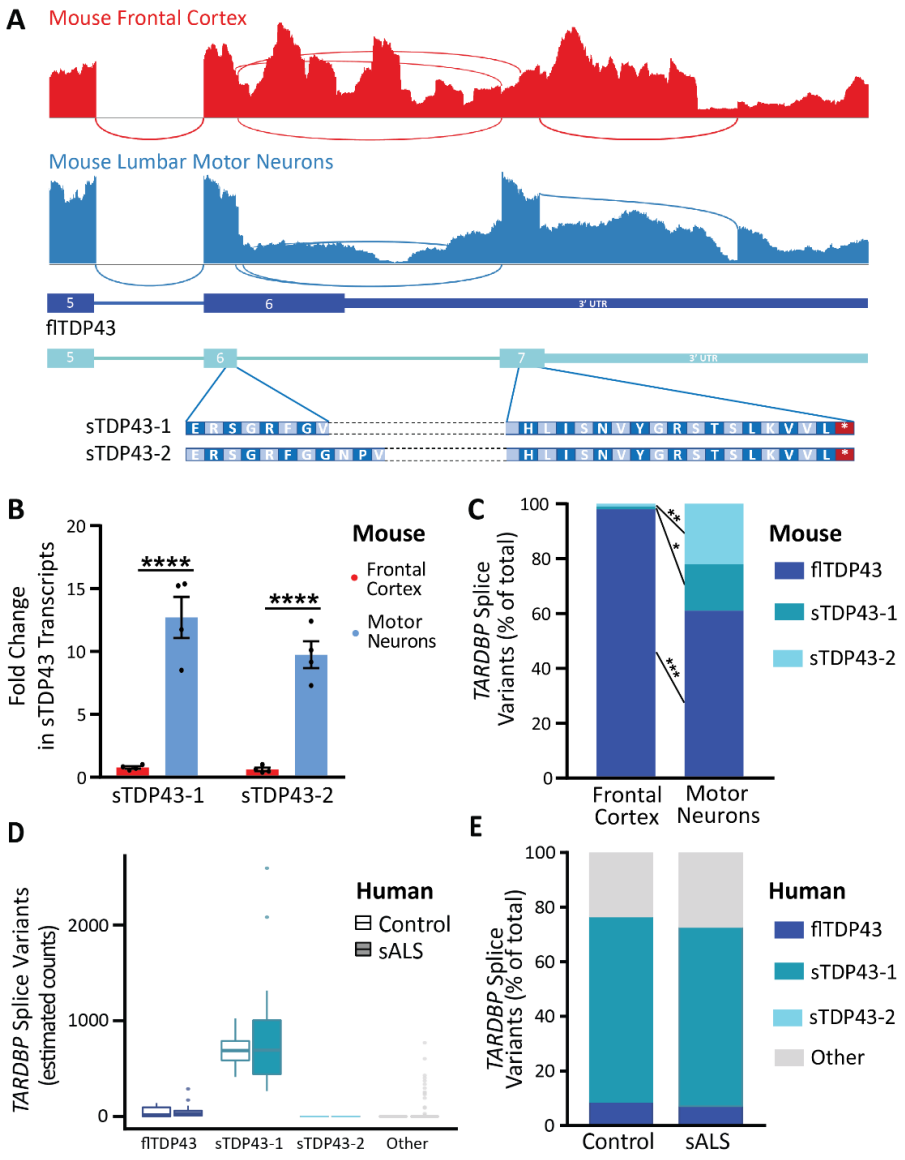


Figure 3.7. sTDP43 transcripts are enriched in lumbar motor neurons. (A) Sashimi plots depicting *TARDBP* splicing in murine frontal cortex homogenate (red) or microdissected lumbar motor neurons (blue). (B) Both sTDP43-1 and sTDP43-2 splice events are highly enriched in lumbar motor neurons compared to frontal cortex homogenate. Graph depicts read counts normalized to reads per million for each library (4 replicates, ****p<0.0001 multiple t-test with the Holm-Sidak correction). (C) While sTDP43-1 and -2 each comprise ~1% of the total *TARDBP* transcripts in frontal cortex homogenate, they make up 17 and 22% of total *TARDBP* transcripts in lumbar motor neurons, respectively (frontal cortex n=6, lumbar motor neurons n=4, *p<0.05, **p<0.01, ***p<0.001, two-way ANOVA with Sidak's multiple comparison test). (D) sTDP43-1 is enriched within lumbar motor neurons microdissected from both control (n=9) and sALS (n=13) patient tissue. (E) sTDP43-1 makes up the majority of total *TARDBP* transcripts in both control and sALS patient lumbar motor neurons.

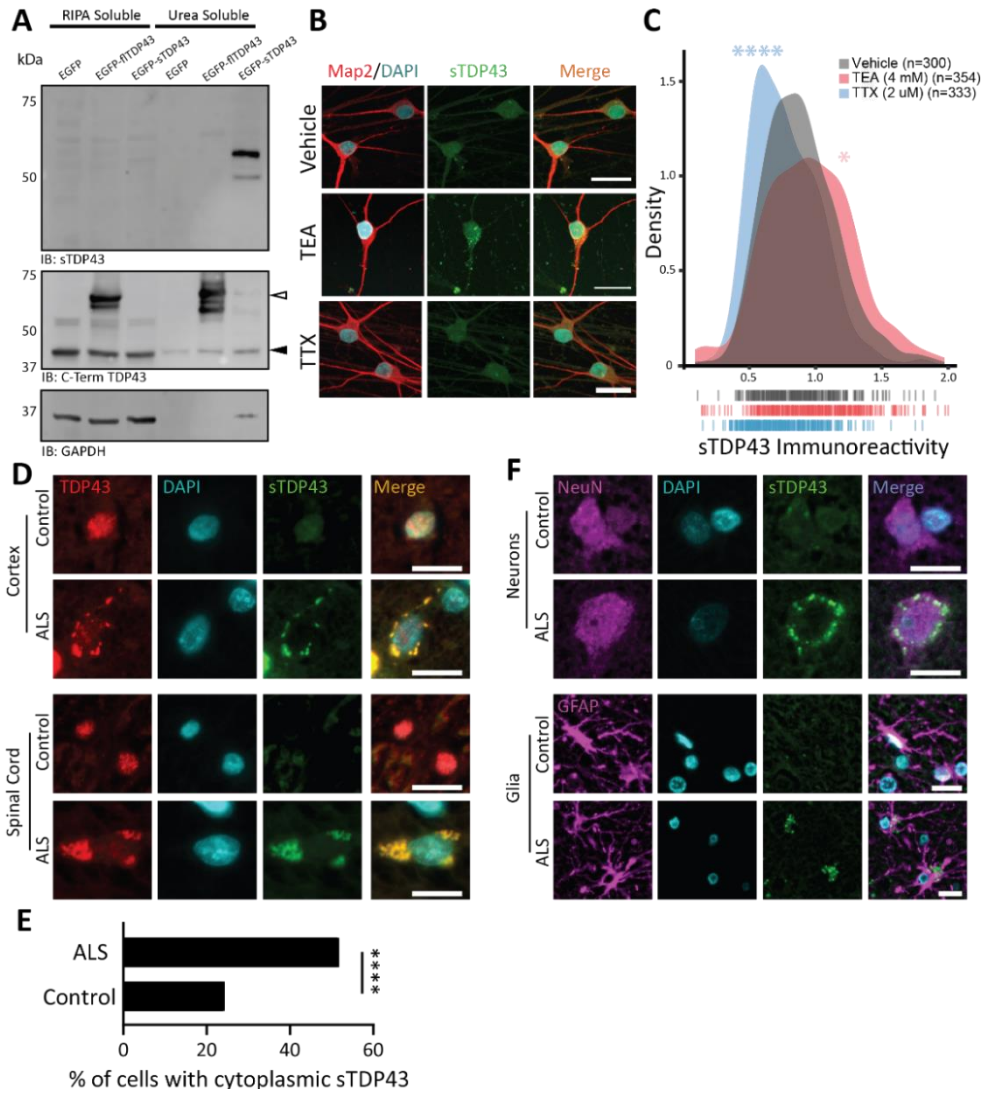
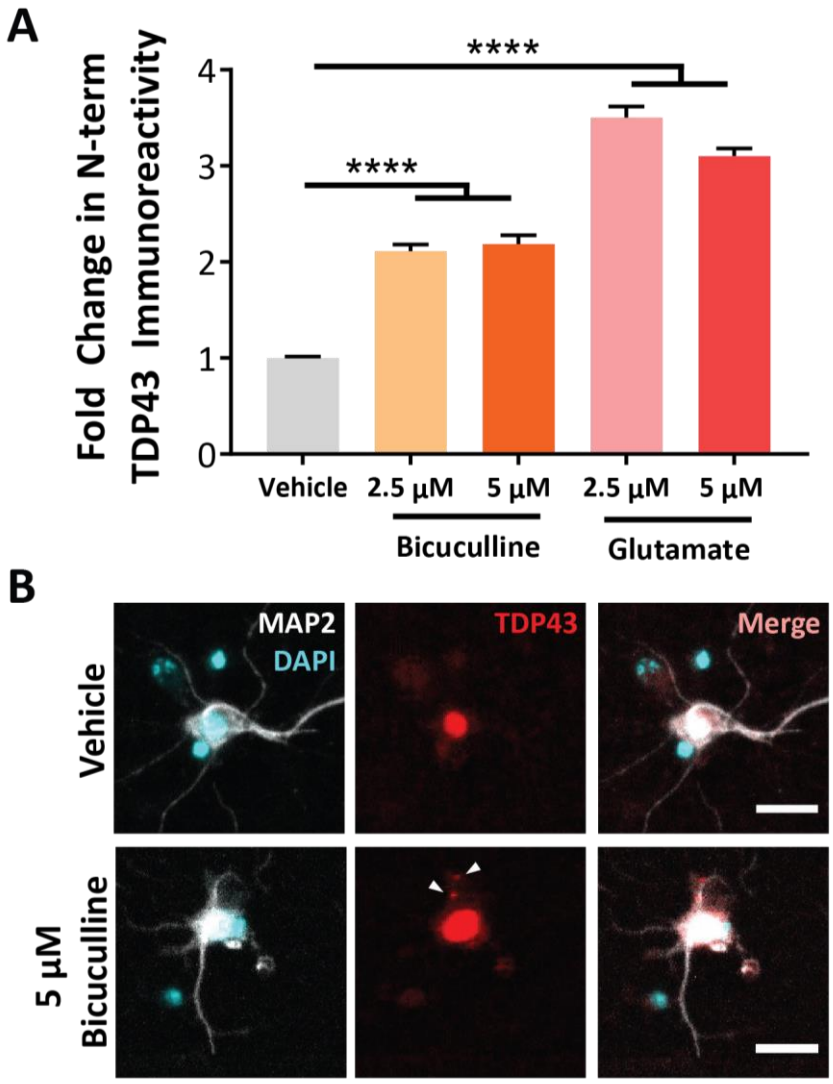
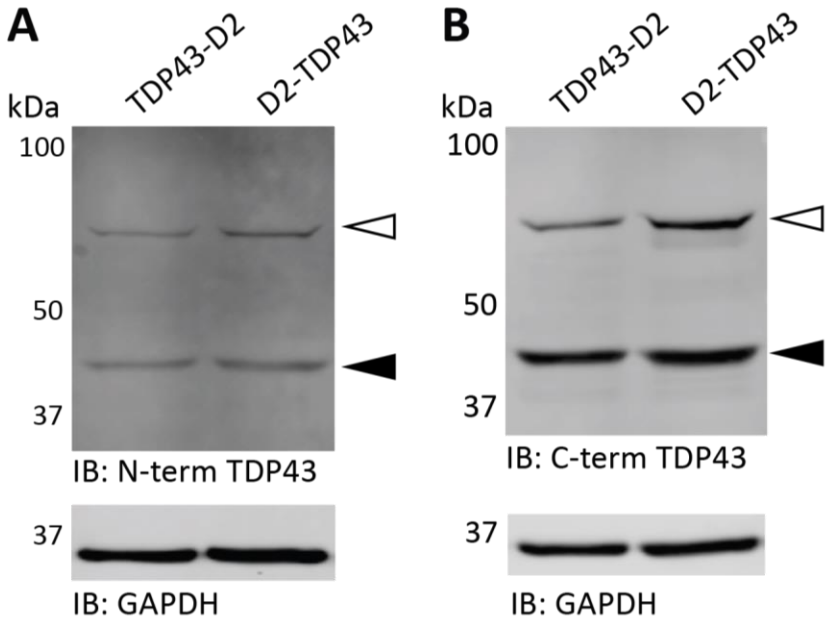


Figure 3.8. Endogenous sTDP43 is detectable *in vivo* by antibodies generated against its novel C-terminus. (A) Western blot of EGFP-tagged flTDP43 or sTDP43 overexpressed in HEK293T cells, demonstrating the insolubility of sTDP43 in RIPA buffer. Black arrowhead, endogenous TDP43; white arrowhead EGFP-flTDP43. (B) ICC using sTDP43 antibodies showed increased immunoreactivity in TEA-treated iNeurons and decreased immunoreactivity in TTX-treated iNeurons. (C) Density plot depicting the change in sTDP43 immunoreactivity between conditions. Vehicle n=300, TEA n=354, TTX n=333, 3 replicates, dashes indicate single neurons, * $p<0.05$, **** $p<0.0001$, Kolmogorov-Smirnov test. (D) IHC comparing the distribution of N-terminal TDP43 and sTDP43 in spinal cord and cortex from patients with sporadic (s)ALS and controls. (E) Quantification of cells with cytoplasmic sTDP43 in control and ALS patient spinal cord (control n=115, ALS n=110, data representative of two control and three ALS patients, ****$p<0.0001$, Fisher's exact test). (F) IHC demonstrating neuronal and glial sTDP43 accumulation in cortex from individuals with sALS and controls. Scale bars in (B), (D), and (F) 20 µm.

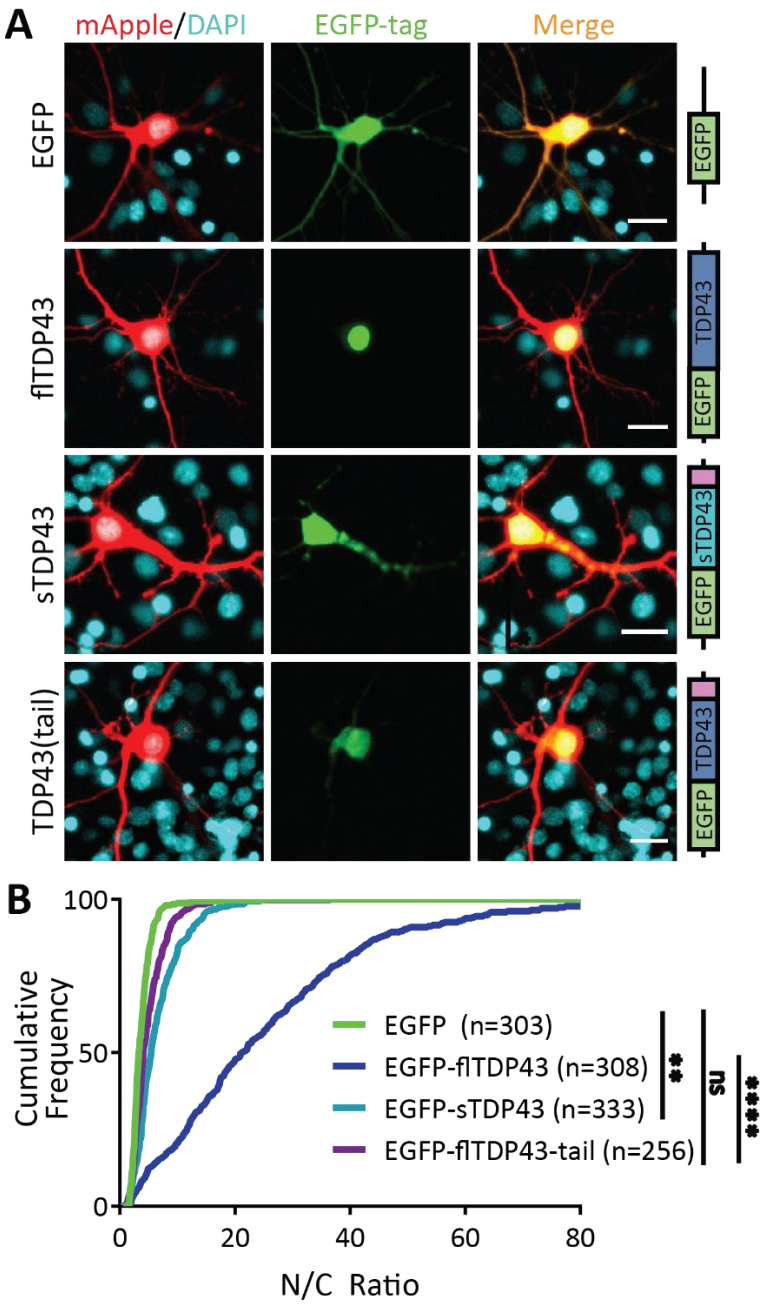
Supplemental Figures



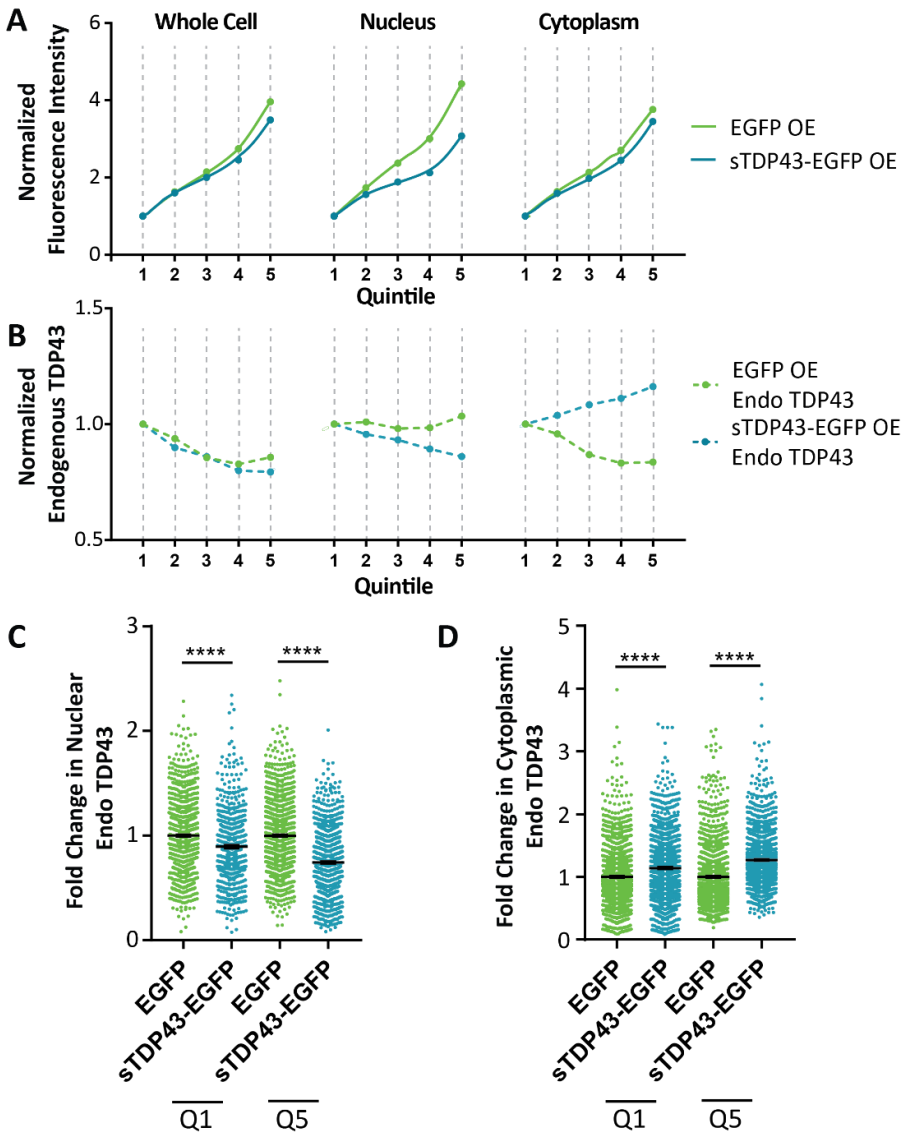
Supplemental Figure 3.1. Multiple drivers of neuronal hyperexcitability upregulate N-terminal TDP43. (A) DIV 28 rat primary mixed cortical neurons treated with glutamate or bicuculline for 48h show an increase in N-terminal TDP43 immunoreactivity compared to vehicle-treated controls (Vehicle n=879, 2.5 µM bicuculline n=1166, 5 µM bicuculline n=837, 2.5 µM glutamate n=1315, 5 µM bicuculline n=1536, data represent two replicates, **** $p<0.0001$, one-way ANOVA with Dunnett's post-test). (B) Representative images of TDP43 staining in vehicle- or bicuculline-treated neurons. White arrows indicate cytosolic puncta. Scale bar in (B), 20 µM.



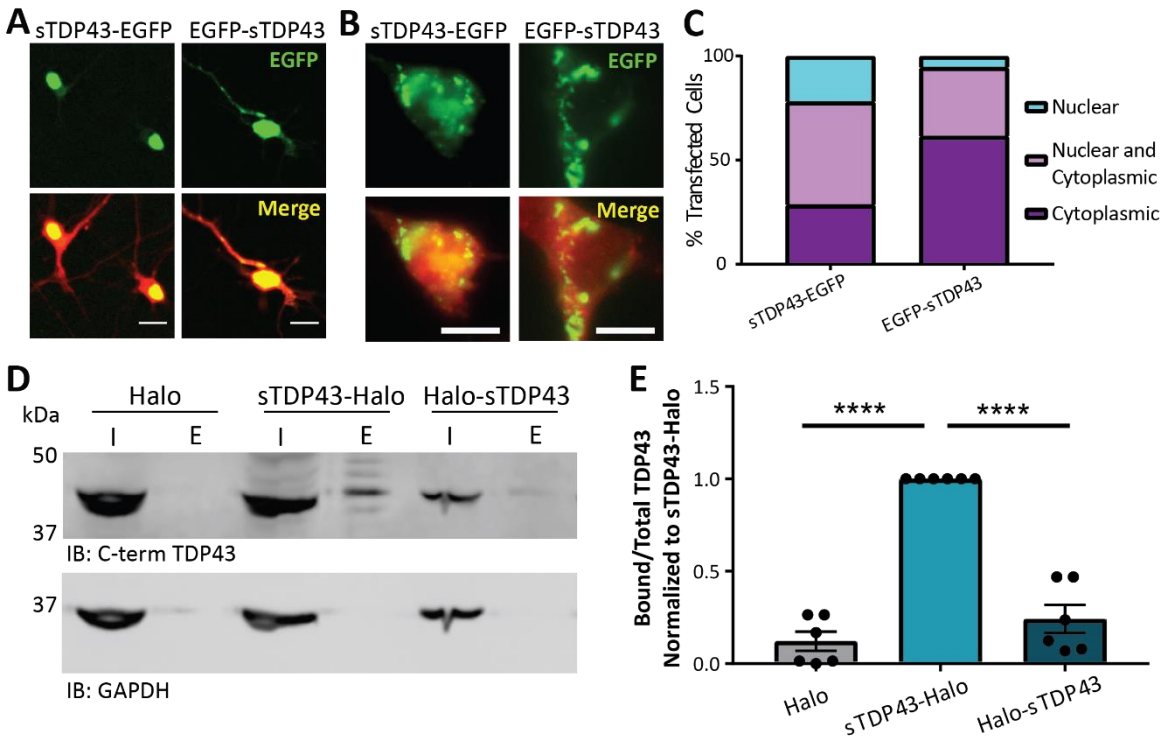
Supplemental Figure 3.2. Validation of Dendra2-tagged iPSC lines. iPSCs immunoblotted for both (A) N-terminal and (B) C-terminal TDP43 indicate that both iPSCs lines are heterozygous for the insertion of Dendra2 (white arrow). Black arrow indicates untagged TDP43, GAPDH served as a loading control.



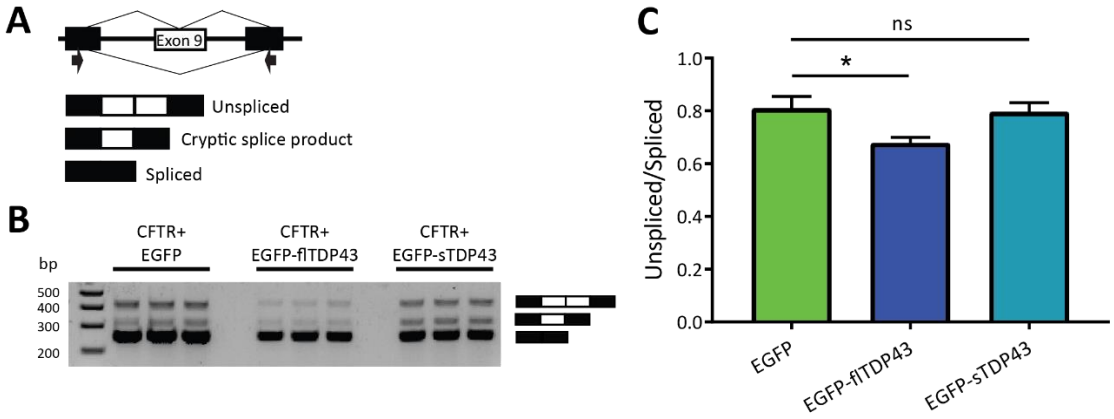
Supplemental Figure 3.3. The sTDP43 C-terminal tail shifts flTDP43 localization to the cytoplasm. (A) Rat primary mixed cortical neurons were transfected with mApple and EGFP-tagged TDP43 isoforms, then imaged by fluorescence microscopy. (B) Much like EGFP-sTDP43, EGFP-flTDP43-tail is significantly more cytosolic than flTDP43-EGFP (EGFP n=303, EGFP-flTDP43 n=308, EGFP-sTDP43 n=333, EGFP-flTDP43-tail n=256, consistent among 3 replicates, **p<0.01, ****p<0.0001, one-way ANOVA with Dunnett's post-test). Scale bar, 20 µm.



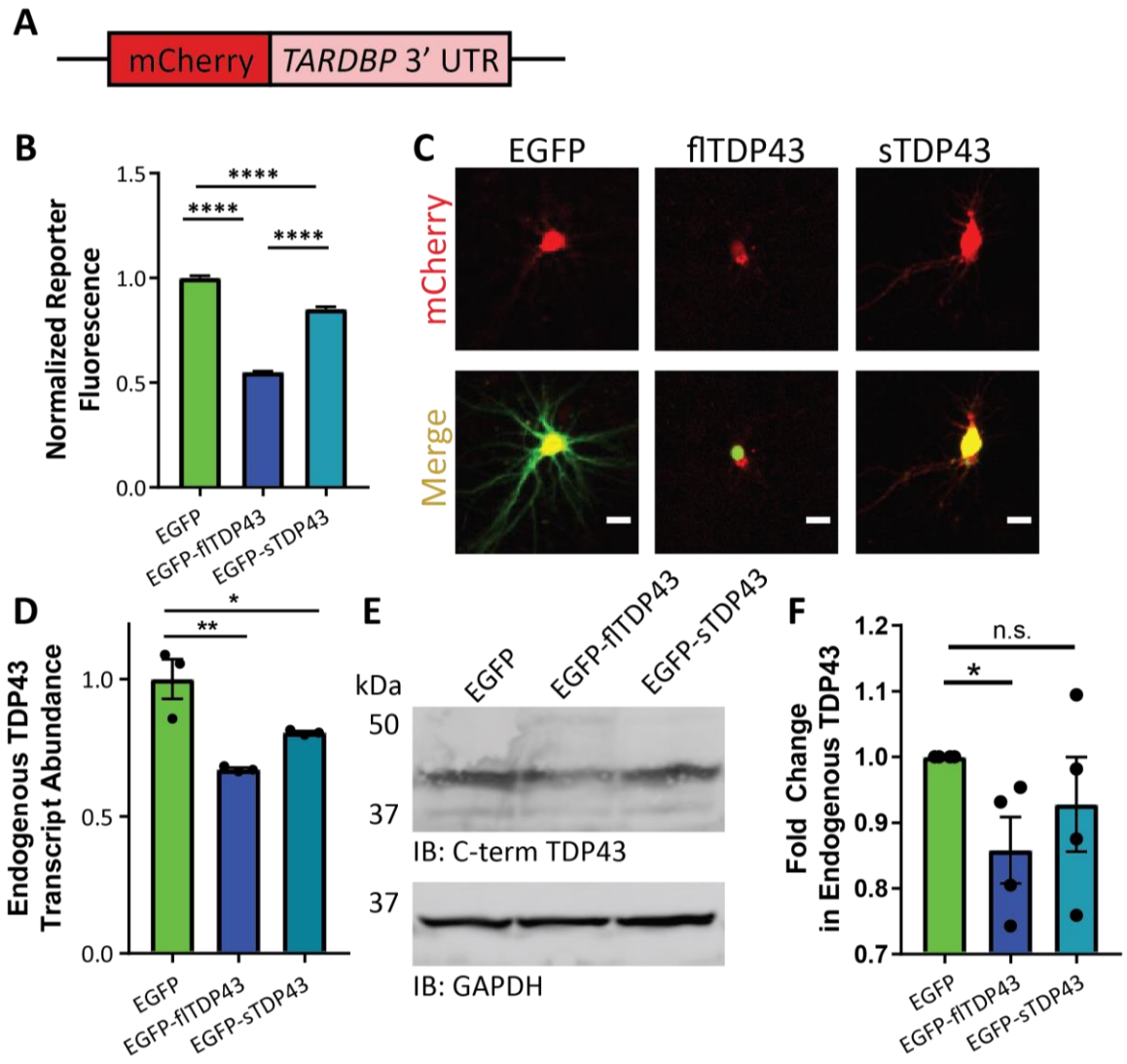
Supplemental Figure 3.4. sTDP43 drives endogenous TDP43 mislocalization in a dose-dependent manner. (A) HEK293T cells were transfected with EGFP or EGFP-tagged sTDP43, then immunostained using an antibody that recognizes the C-terminus of endogenous (endo) TDP43. Whole cell ROIs were determined by mApple fluorescence, and nuclear ROIs were determined by DAPI staining. Cells were sorted into quintiles based on EGFP fluorescence. (B) Endo TDP43 signal is comparable between cells overexpressing EGFP and sTDP43-EGFP at the whole cell level. However, nuclear endo TDP43 is reduced in a dose-dependent manner in cells overexpressing sTDP43-EGFP while cytosolic endo TDP43 increases in a dose-dependent manner. (C) Both low (Q1) and high (Q5) expression of sTDP43-EGFP results in a reduction in nuclear endo TDP43 (Q1 EGFP n=1117, Q1 sTDP43-EGFP n=638, Q5 EGFP n=1203, Q5 n=1005, ****p<0.0001, two-tailed t-test). (D) Cytoplasmic endogenous TDP43 is elevated by the expression of sTDP43-EGFP in both low (Q1) and high (Q5) expressing cells (Q1 EGFP n=1599, Q1 sTDP43-EGFP n=1599, Q5 EGFP n=1419, Q5 sTDP43-EGFP n=1492, ****p<0.0001, two-tailed t-test).



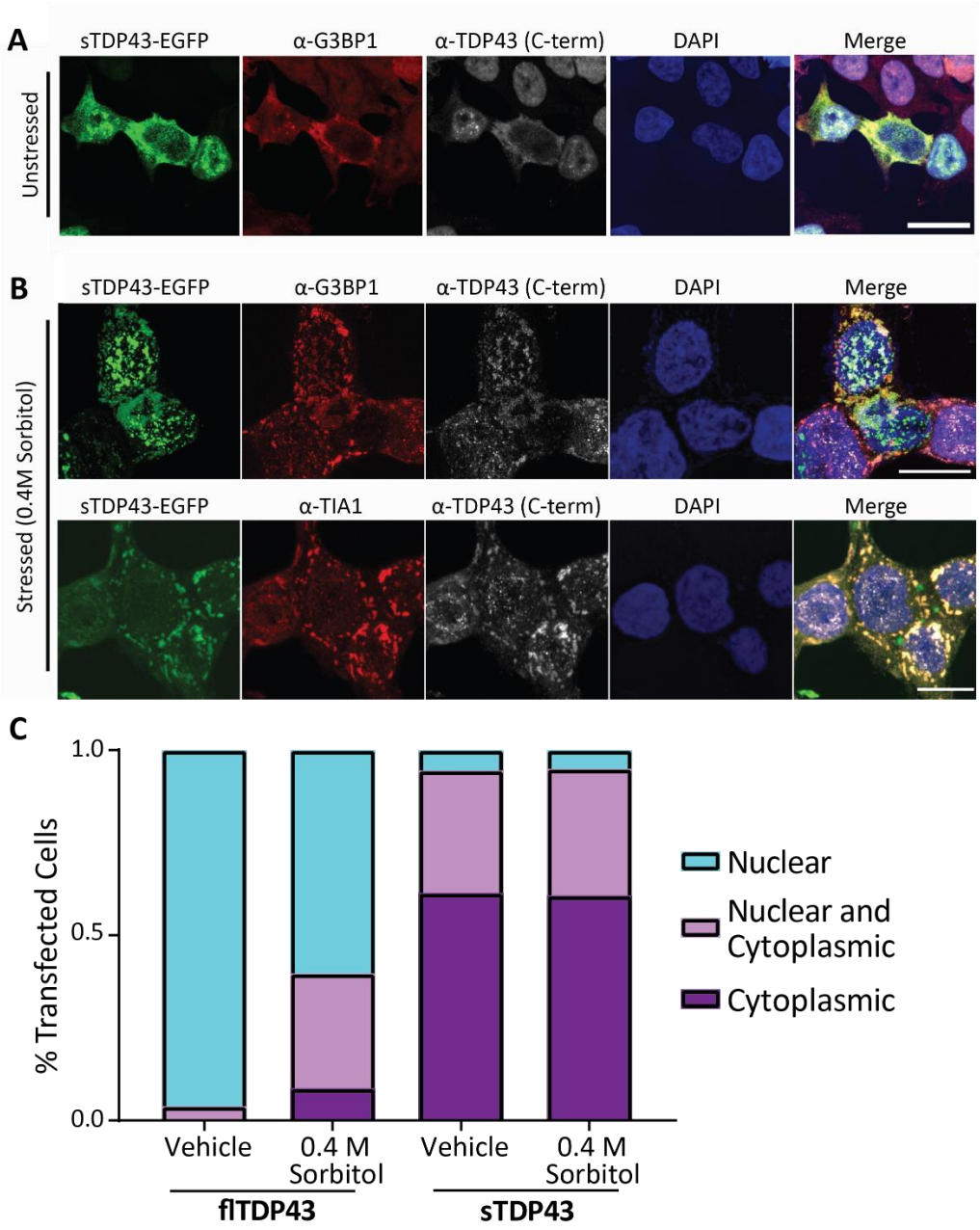
Supplemental Figure 3.5. Position of exogenous protein tags influence sTDP43 localization and binding. (A) Fluorescence microscopy of rodent primary mixed cortical neurons or (B) HEK293T cells expressing N- or C-terminally tagged sTDP43 (green) as well as the whole cell marker mApple (red). (C) Characterization of sTDP43 localization in HEK293T cells expressing C- or N-terminally tagged sTDP43 (sTDP43-EGFP n=196, EGFP-sTDP43 n=200). (D) sTDP43 fused with an N- or C-terminal HaloTag was expressed in HEK293T cells and immunoprecipitated with HaloLink resin. Bound, endogenous TDP43 was immunoblotted with a C-terminal TDP43 antibody. GAPDH served as a loading control. Input, (I); eluate, (E). (E) Quantification of data shown in (D), demonstrating the fraction of total TDP43 bound to HaloTag-sTDP43, sTDP43-HaloTag, or HaloTag alone. Data were combined from 5 replicates, ****$p<0.0001$, one-way ANOVA with Dunnett's post-test.quantified in (E). Scale bars in (A) 20 µm, (B) 10 µm.



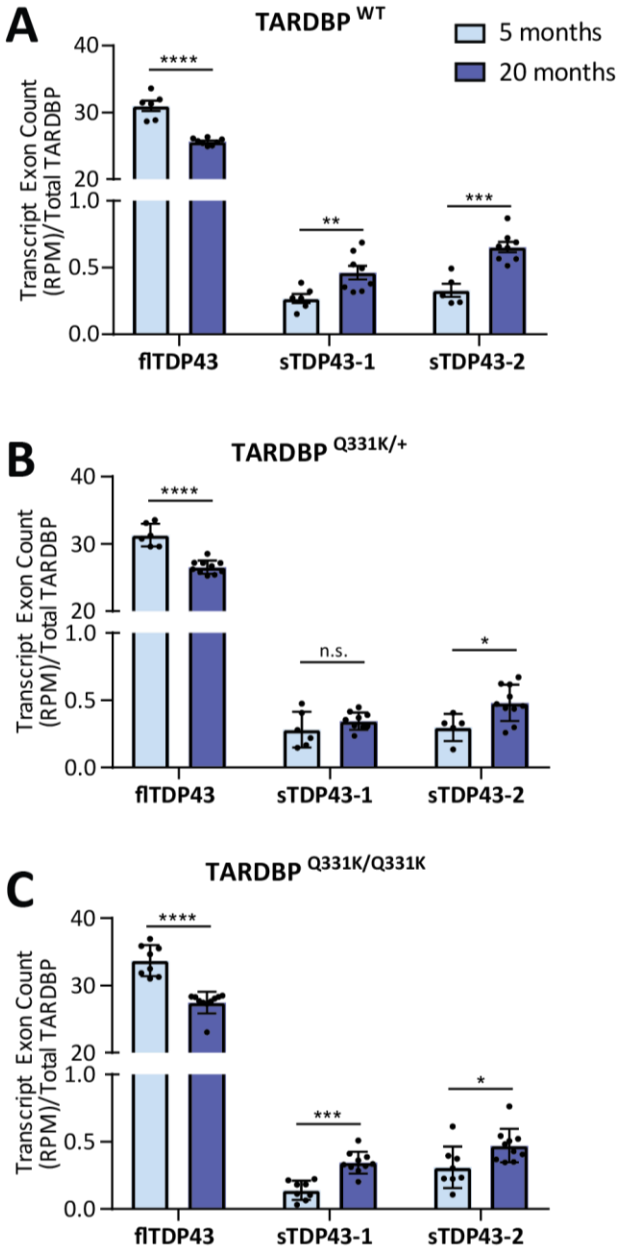
Supplemental Figure 3.6. sTDP43 is deficient in splicing activity. (A) Schematic of the *CFTR* minigene reporter. TDP43-mediated splicing of the reporter results in exon 9 exclusion. Arrows indicate primers used to amplify the splice junction. (B-C) EGFP-flTDP43, but not EGFP-sTDP43, effectively excludes *CFTR* exon 9 in HEK293T cells overexpressing the reporter (3 replicates, *p<0.05, one-way ANOVA, Dunnett's post-test).



Supplemental Figure 3.7. sTDP43 autoregulatory function is impaired. (A) Schematic of the TDP43 autoregulation reporter in which the fluorescent protein mCherry is fused to the *TARDBP* 3' UTR. (B) Rodent primary mixed cortical neurons overexpressing EGFP-flTDP43 show a significant reduction in reporter signal, consistent with autoregulation, while those overexpressing EGFP-sTDP43 do so to a lesser degree (EGFP n=2044, EGFP-flTDP43 n=2375, EGFP-sTDP43 n=2208, 3 replicates, ****p<0.0001, one-way ANOVA, Dunnett's post-test). (C) Fluorescence microscopy of rodent primary mixed cortical neurons demonstrating reduced reporter fluorescence in neurons co-expressing EGFP-flTDP43, in comparison to those co-expressing EGFP or EGFP-sTDP43. (D) qRT-PCR of HEK293T cells overexpressing EGFP, EGFP-flTDP43, or EGFP-sTDP43. Endogenous full-length *TARDBP* transcript was detected using primers that flank the stop codon, and transcript levels were normalized to GAPDH. *p<0.05, **p<0.01, one-way ANOVA with Dunnett's post-test. (E) HEK293T cells overexpressing each construct were immunoblotted for C-terminal TDP43, with GAPDH serving as a loading control. Results quantified in (F) (4 replicates, * p<0.05, one-way ANOVA, Dunnet's post-test). Scale bar in (C), 10 µm.



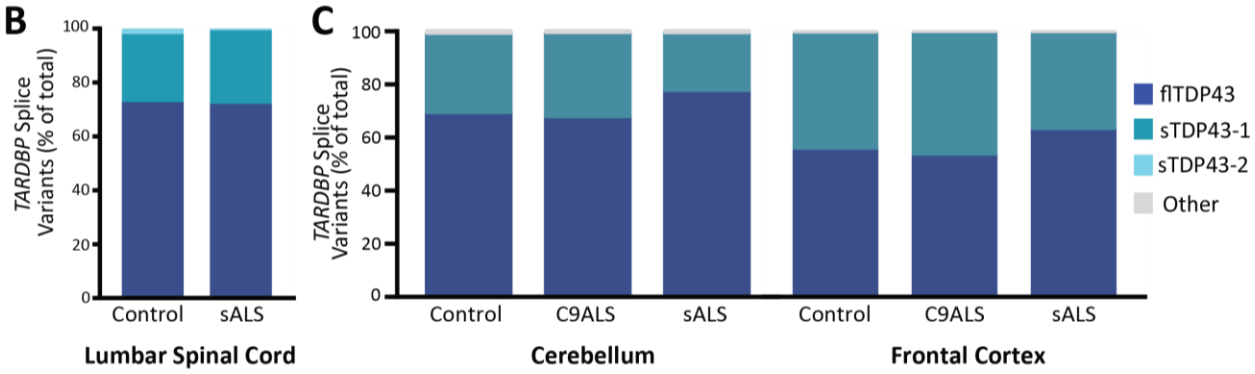
Supplemental Figure 3.8. sTDP43 colocalizes with components of stress granules. (A) HEK293T cells were transfected with EGFP-tagged sTDP43 and immunostained using antibodies against the stress granule marker G3BP1 and endogenous TDP43. (B) Overexpressed sTDP43-EGFP colocalizes with endogenous TDP43 and stress granule markers G3BP1 and TIA1 in HEK293T cells treated with 0.4M sorbitol. (C) Assessment of stress-dependent changes in flTDP43 and sTDP43 localization. HEK293T cells were transfected with each construct and then stressed with sorbitol as before. The distribution of each protein was characterized as primarily nuclear, primarily cytoplasmic, or both for each cell (flTDP43 + vehicle n=216, flTDP43 + sorbitol n=219, sTDP43 + vehicle n=200, sTDP43 + sorbitol n=158). Scale bar in (A) 20 µm, scale bar in (B), 10 µm.



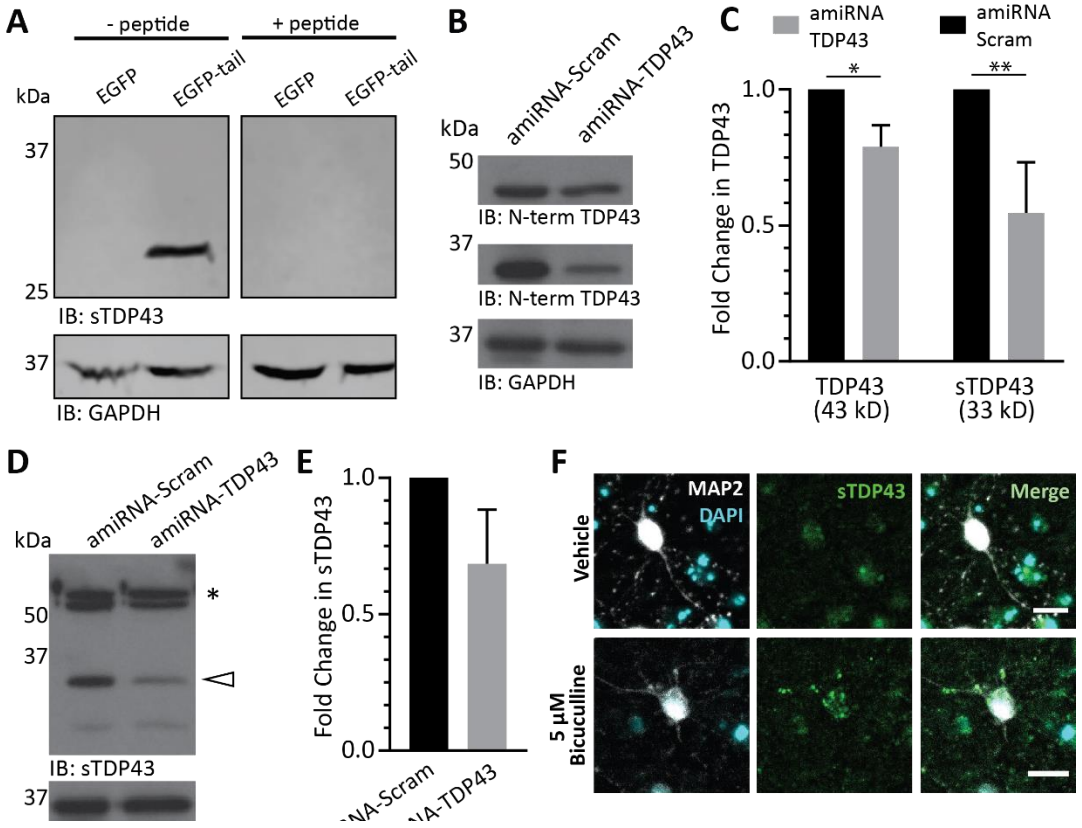
Supplemental Figure 3.9. sTDP43 transcript abundance increases with age. Murine frontal cortex collected at 5 (pale blue) or 20 (purple) months of age show both a decrease in flTDP43 and an increase in sTDP43-1 and -2 transcript abundance with age. This is observed in WT mice (A) as well as those that are hetero- (B) and homozygous (C) for *TARDBP*(Q331K) mutations. Graph depicts read counts normalized to reads per million for each library as a fraction of total *TARDBP* (4 replicates, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 multiple t-test with the Holm-Sidak correction).

A

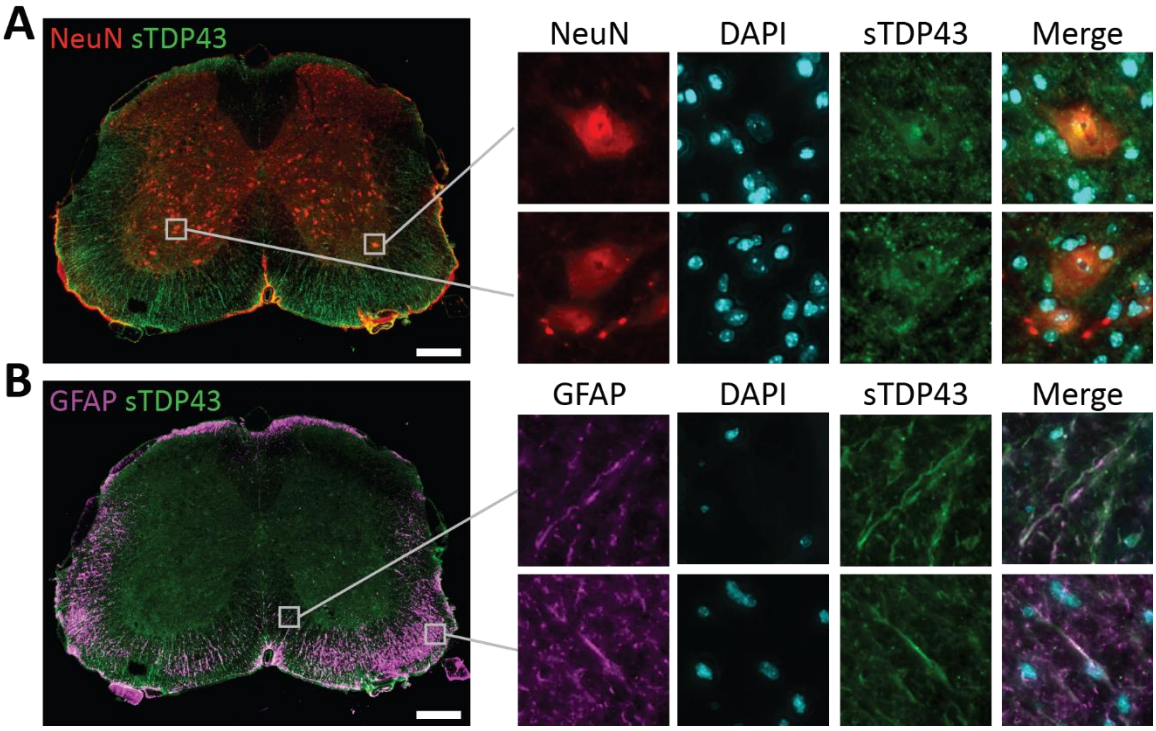
Original Study	Species	Library preparation	Read Length	Cell or Tissue Type	Disease Condition	n
White et al. 2018	Mouse	Clontech SMART-seq and Illumina Nextera X	100 bp, paired	Laser captured spinal motor neuron	Wildtype	4
White et al. 2018	Mouse	Illumina TruSeq	100 bp, paired	Frontal cortex	Wildtype	6
Krach et al. 2018	Human	NuGEN Ovation	50 bp, single	Laser captured spinal motor neuron	Control and sALS	21
D'Erchia et al. 2017	Human	Illumina TruSeq	100 bp, paired	Ventral horn spinal cord	Control and sALS	11
Preduncio et al. 2015	Human	Illumina TruSeq	100 bp, paired	Cerebellum and Frontal cortex	Control, C9ALS, and sALS	53



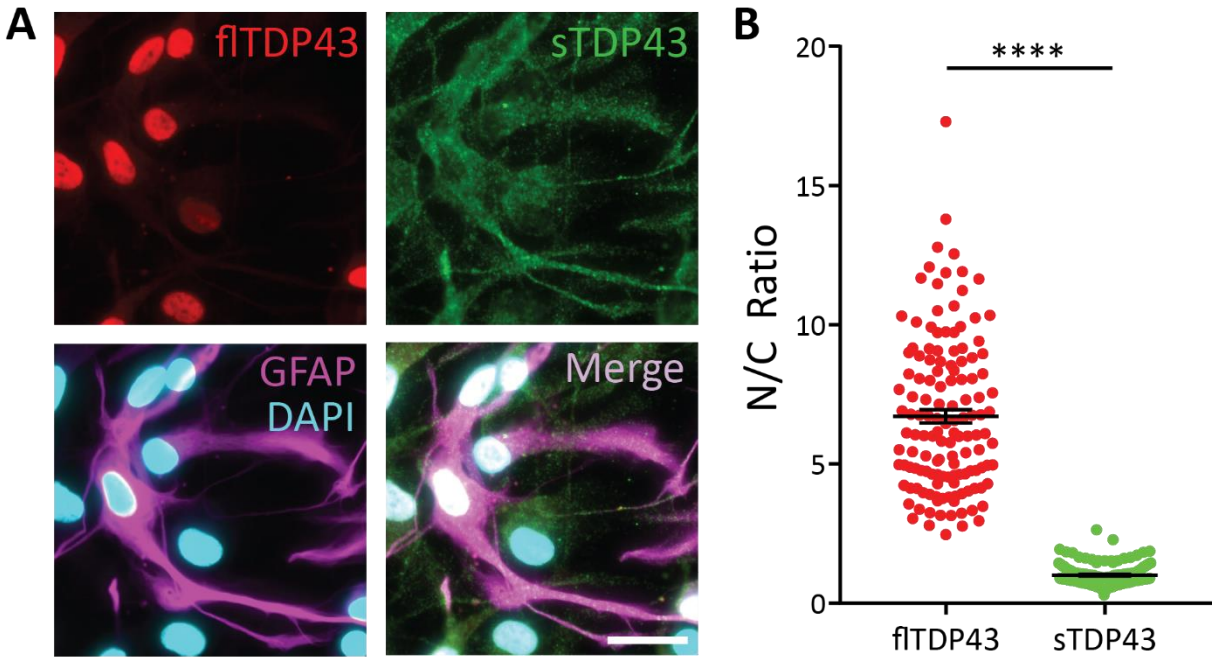
Supplemental Figure 3.10. sTDP43 transcripts are present in a variety of tissue types and disease states. (A) Summary of previously published RNA-seq studies analyzed for sTDP43. (B) sTDP43-1 represents 30% of *TARDBP* transcripts in spinal cord ventral horn homogenate isolated from both control and sALS patients, with the remainder corresponding to flTDP43. (C) Similarly, sTDP43-1 makes up 30% of *TARDBP* transcripts in cerebellum, and 55% of *TARDBP* transcripts in frontal cortex. In each case, sTDP43-2 transcripts were largely undetectable, and there was no significant change in transcript isoform abundance in C9ALS or sALS patients compared to controls.



Supplemental Figure 3.11. Validation of the sTDP43-specific antibody. (A) The sTDP43 antibody specifically recognizes EGFP fused to the 18-amino acid C-terminus of sTDP43, but not EGFP alone. Preincubation with a peptide corresponding to the sTDP43 C-terminal tail abolishes the signal. (B) N2A cells were transiently transfected with artificial microRNA (amiRNAs) targeting TDP43, and proteins were separated by SDS-PAGE and immunoblotted using antibodies against N-terminal TDP43. Two bands were detected, the first at 43 kD corresponding to flTDP43 and the second at 33 kD corresponding to sTDP43. GAPDH served as a loading control. (C) Compared to scrambled amiRNA, cells expressing amiRNA-TDP43 show a ~30% reduction of the 43kD species and a ~65% reduction of the 33kd species (3 replicates, *p<0.05, **p<0.01, two-tailed t-test). (D) Immunoblotting with the sTDP43-specific antibody detects a 33 kD band (white arrow), as well as a non-specific band at ~55 kD (asterisk). (E) sTDP43 shows a ~40% reduction in cells expressing amiRNA-TDP43 compared to scrambled control. (F) DIV 28 rodent primary mixed cortical neurons were treated with 5 uM bicuculline for 48h and immunostained for sTDP43. Bicuculline-treated neurons show cytosolic sTDP43 inclusions that are absent in vehicle-treated controls. Scale bar in (F), 20 µM.



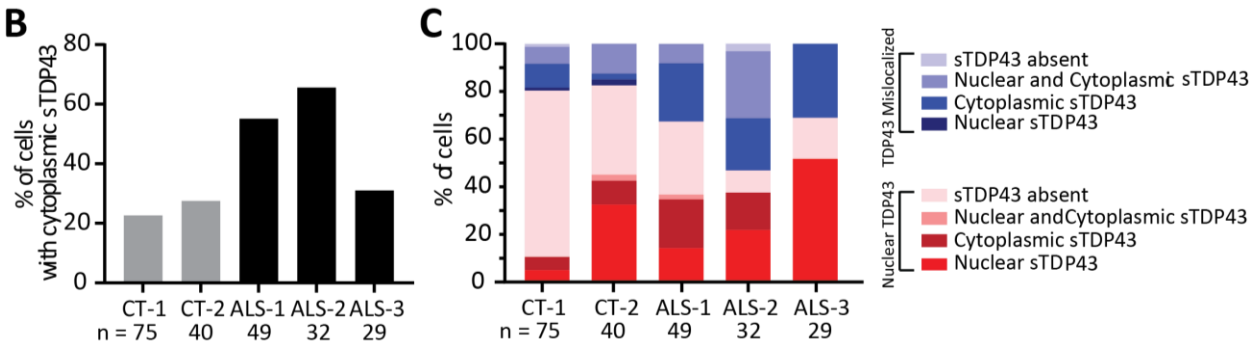
Supplemental Figure 3.12. Endogenous sTDP43 is expressed by neurons and glia in murine lumbar spinal cord. Immunohistochemistry in murine spinal cord, showing colocalization of sTDP43 immunoreactivity (green) with the neuronal marker NeuN (red) in the ventral horn (A) and with the astrocytic marker GFAP (B, purple). Scale bars (A) and (B), 200 µm.



Supplemental Figure 3.13. Endogenous sTDP43 is produced by human iPSC-derived astrocytes. (A) Immunocytochemistry using antibodies against flTDP43 (red) and sTDP43 (green) in astrocytes differentiated from human iPSCs. (B) Reflecting their unique subcellular distributions, flTDP43 displays a significantly higher nuclear-cytoplasmic ratio (NCR) than sTDP43 (3 replicates, flTDP43 n=136, sTDP43 n=136, ****p<0.0001, two-tailed t-test). Scale bar in (A) 10 µm.

A

Case	Sex	Age	Age of Onset	PMI (Hours)	Clinical Diagnosis	Neuropath Diagnosis
CT-1	Male	56	49	14	MSA-P	MSA
CT-2	Male	88	NA	5	Normal	DLB with AD
ALS-1	Female	55	53	Unknown	Bulbar Onset ALS	C9 ALS
ALS-2	Female	60	Unknown	33	Limb Onset ALS	sALS
ALS-3	Female	87	77	5	Bulbar Onset ALS	sALS



Supplemental Figure 3.14. Characterization of sTDP43 pathology in ALS-patient tissue. (A) Additional information for the control and ALS patient tissue used in these studies. PMI, post-mortem interval; MSA-P, multiple system atrophy with Parkinsonism; DLB with AD, Dementia with Lewy bodies with concurrent Alzheimer's Disease. (B) Percentage of cells with cytoplasmic sTDP43 identified in each control (gray) or ALS (black) patient. Number of cells counted per sample is listed below each column. (C) Characterization of sTDP43 localization for each patient. These data are further divided based on whether flTDP43 was nuclear (red) or cytoplasmically mislocalized (blue).

Species	Amino Acid Sequence																	
Reference Sequence	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Chimpanzee (*Pan troglodytes*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Orangutan (*Pongo abelii*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Baboon (*Papio anubis*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Drill (*Mandrillus leucophaeus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Black and white colobus (*Colobus angolensis palliatus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Goat (*Capra hircus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Cat (*Felis catus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Alpine marmot (*Marmota marmota marmota*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Long-tailed chincilla (*Chinchilla lanigera*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Ground squirrel (*Ictidomys tridecemlineatus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Kangaroo rat (*Dipodomys ordii*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Guinea pig (*Cavia porcellus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Brown rat (*Rattus norvegicus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Darama mole rat (*Fukomys damarensis*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Blind Mole Rat (*Nannospalax galili*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Golden hampster (*Mesocricetus auratus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Deer mouse (*Peromyscus maniculatus bairdii*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
House mouse (*Mus musculus*)	V	H	L	I	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Naked mole rat (*Heterocephalus glaber*)	V	H	L	I	S	N	V	F	G	R	S	T	S	L	K	V	V	L
Horse (*Equus calabus*)	V	H	L	M	S	N	V	Y	G	R	S	T	S	L	K	V	V	L
Tibetan antelope (*Pantholops hodgsonii*)	V	H	L	I	S	N	V	H	G	R	S	T	S	L	K	V	V	L
Chinese tree shrew (*Tupaia chinensis*)	V	H	L	I	S	N	V	S	G	R	S	T	S	L	K	V	V	L

Supplemental Table 3.1. Amino acid sequence of the sTDP43 C-terminal tail is highly conserved.

Species	Isoform 1 Identity	Isoform 2 Identity
Baboon (Papio anubis)	99%	99%
Drill (Mandrillus leucophaeus)	96%	96%
Angolan black and white colobus (Colobus angolensis palliatus)	96%	96%
Goats (Capra hircus)	96%	96%
Chinese tree shrew (Tupaia chinensis)	96%	96%
Horse (Equus caballus)	96%	96%
Cat (Felus catus)	96%	96%
Ground Squirrel (Ictidomys tridecemlineatus)	96%	96%
Shrew (Mus pahari)	96%	96%
Brown rat (Rattus novegicus)	96%	96%
Deer mouse (Peromyscus maniculatus bairdii)	96%	96%
Prairie vole (Microtus ochrogaster)	96%	96%
House mouse (Mus musculus)	96%	96%
Ryukyu mouse (Mus caroli)	96%	96%
Damaraland mole rat (Fukomys damarensis)	96%	96%
Tibetan antelope (Pantholops hogsonii)	96%	96%
Blind mole rat (nannospalax galili)	96%	96%
Chinese hamster (Cricetulus griseus)	95%	95%
Guinea pig (Cavia procellus)	95%	95%
Golden hampster (Mesocricetus auratus)	95%	92%
Lesser Egyptian jerboa (Jaculus jaculus)	95%	93%
Cow (Bos tauru)	95%	95%

Supplemental Table 3.2. Nucleotide sequence of the sTDP43-1 and -2 splice junctions are highly conserved.

Construct	Source	Complementary Oligomers	Sequence (5' to 3')
pUCM-CLYBL-NGN1-2-RFP	Gift from M. Ward		
pLTC13-L1	Gift from M. Ward		
pLTC13-R1	Gift from M. Ward		
pX335-U6-Chimeric_BB-CBh-hSpCas9n(D10A)	Addgene (42335, donated by Feng Zhang)		
pX330S-4	Addgene (58780, donated by Feng Zhang)		
pUCM-N-term-TARDBP-D2-HDR	Synthesized by Blue Heron, LLC		
pX335-sgRNA-D2-TDP43-Upstream	This paper	Sense	GACCGTTCATATCTCTTTTCTCTTT
		Antisense	AAACAAAGAGAAAAGAGATATGAAC
pX335-sgRNA-D2-TDP43-Downstream	This paper	Sense	CACCGGGGCTCATCGTTCTCATCTT
		Antisense	AAACAAGATGAGAACGATGAGCCCC
pUCM-C-term-TARDBP-D2-HDR	Synthesized by Blue Heron, LLC		
pX335-sgRNA-TDP43-D2-Upstream	This paper	Sense	CACCGGTTGGTTGGTATAGAATGG
		Antisense	AAACCCATTCTATACCAACCAACC
pX335-sgRNA-TDP43-D2-Downstream	This paper	Sense	CACCGACCACTGCCCGACCCTGCAT
		Antisense	AAACATGCAGGGTCGGGCAGTGGTC

Supplemental Table 3.3. Constructs and primer sequences used to generate iPSC lines.

Supplemental Table 4. Primary antibodies

Antibody	Source	Catalog Number	Species	Dilution
Vglut1	Synaptic Systems	135303	rabbit	1:200 for ICC in iNeurons
Tuj1	BioLegend	801202	mouse	1:500 for ICC in iNeurons
N-term TDP43	Sephton et al. 2011; Barmada et al. 2014	NA	rabbit	1:5000 for ICC and western
C-term TDP43	Sephton et al. 2011; Barmada et al. 2014	NA	rabbit	1:5000 for ICC and western
sTDP43	Custom-made from Genscript	NA	rabbit	1:1000 for ICC and western
				1:500 for murine IHC
				1:250 for human IHC
Map2	Milipore	MAB3418	mouse	1:1000 for ICC in iNeurons
GAPDH	Milipore	MAB374	mouse	1:1000 for ICC in iNeurons
GFAP	Abcam	AB53554	goat	1:500 for murine IHC
NeuN	Abcam	AB104225	mouse	1:500 for murine IHC
GFAP	Milipore	AB5541	chicken	1:500 for human IHC
NeuN	Milipore	MAB377	mouse	1:250 for human IHC
TDP43	R&D Biosystems	MAB7778	mouse	1:1000 for western and ICC
				1:500 for ICC in iPSC-derived astrocytes
				1:250 for human IHC
GFAP	Sigma	C9205	mouse	1:500 for ICC in iPSC-derived astrocytes

Supplemental Table 3.4. Primary antibodies.

Supplemental Table 5. Primers used in qRT-PCR and CFTR splicing assays

Target	Location	Primer	Sequence (5' to 3')
ARC	*ARC* exon 2	Forward	CCTGTACCAGACGCTCTACG
		Reverse	GCAGGAAACGCTTGAGCTTG
Total TARDBP	*TARDBP* exon 1 and 2	Forward	CTGCTTCGGTGTCCCTGTC
		Reverse	TGGGCTCATCGTTCTCATCT
full-length (fl) TARDBP	*TARDBP* exon 6 stop codon	Forward	GTGGCTCTAATTCTGGTGCAG
		Reverse	CACAACCCCACTGTCTACATT
sTDP43-1	*sTDP43-1* splice donor	Forward	AGAAGTGGAAGATTTGGTGTTCA
		Reverse	GCATGTAGACAGTATTCCTATGGC
sTDP43-2	*sTDP43-2* splice donor	Forward	AGATTTGGTGGTAATCCAGTTCA
		Reverse	GGCCTGTGATGCGTGATGA
CFTR minigene	*CFTR* exon 9 splice junction	Forward	CAACTTCAAGCTCCTAAGCCACTGC
		Reverse	TAGGATCCGGTCACCAGGAAGTTGGTTAAATCA

Supplemental Table 3.5. Primers used in RT PCR and CFTR assays.

Supplemental Table 6. Source and construction of plasmid vectors

Plasmid	Source	Amplicon or Insert	Primer	Sequence (5' to 3')
pGW1-mApple	Barmada et al. 2014			
pGW1-EGFP(1)	Arrasate et al. 2004			
pGW1-TDP43-EGFP	Barmada et al. 2014			
pGW1-sTDP43-EGFP	This paper	sTDP43	NA	
pGW1-sTDP43(mNES)-EGFP	This paper	sTDP43(mNES)	Forward	CGCGGGCCCATGTCTGAATATATTCG
			Reverse	GCGACCGGTCGCAGCACTCCACCCCCTCCGCCGCTTCTTCCATAAAC
pGW1-EGFP(2)	This paper	EGFP	Forward	GCGAAGCTTGCCACCATGGTGAGCAAG
			Reverse	CGCGGTACCCTTGTACAGCTCGTCCAT
pGW1-EGFP-tail	This paper	tail	Forward	CGTTCATCTCATTTCAAATGTTTATGGAAGAAGCACTTCATTGAAAGTAGTGCTGTAAG
			Reverse	CTAGCTTACAGCACTACTTTCAATGAAGTGCTTCTTCCATAAACATTTGAAATGAGATGAACGGTA
pGW1-EGFP-tail(mNES)	This paper	tail(mNES)	Forward	CGTTCATCTCATTTCAAATGTTTATGGAAGAAGCGGCGGAGGGGGTGGAGTGCTGTAAG
			Reverse	CTAGCTTACAGCACTCCACCCCCTCCGCCGCTTCTTCCATAAACATTTGAAATGAGATGAACGGTA
pGW1-EGFP-flTDP43	This paper	TDP43	Forward	GCG GGT ACC ATG TCT GAA TAT ATT CGG
			Reverse	CGC GCT AGC TTA TCC CCA GCC AGA AG
pGW1-EGFP-sTDP43	This paper	sTDP43	Forward	GCGGGTACCATGTCTGAATATATTCGG
			Reverse	CGCGCTAGCTTACAGCACTACTTTCAATG
pGW1-Halo	This paper	Halo	Forward	AAAAAA TCTAGA GCCACCATGGCAGAAATCGG
			Reverse	AAAAAA CCTGCAGG CTA GGAAATCTCGAGCGTCGACA
pGW1-flTDP43-Halo	This paper	TDP43, Halo	Forward	AAAAAA TCTAGA ATGGCAGAAATCGGTACTGG
			Reverse	AAAAAA CCTGCAGG CTA GGAAATCTCGAGCGTCGACA
pGW1-sTDP43-Halo	This paper	sTDP43	Forward	GCG GCTAGC GCCACC ATGTCTGAATATATTCG
			Reverse	CGCACCGGTGGCAGCACTACTTTC
pCaggs-mCherry-TARDBP3'UTR	This paper	TARDBP Exon 6 and 3'UTR	Forward	ATATTGTACATTGCGCAGTCTCTTTGTGGA
(TDP43 autoregulation reporter)			Reverse	ATATGGCCGAGGCGGCCATCGTGTTTTTCCAGTAAGACTCCAGAC
pcDNA3.1-NLS-mCherry-NES (shuttle-RFP)	Addgene (72660, donated by B. Di Ventura and R. Eils)			
pFN21A HaloTag PUM2 RBD R6SYE	Gift from A. Goldstrohm			
pTB CFTR A455E TG13T5 (CFTR minigene)	Gift from Y. Ayala			
pGW1-CMV	Gift from S. Finkbeiner			
pCaggs-mCherry	Gift from S. Finkbeiner			
pSMART Lenti-shTARDBP (human) CAG-TurboRFP-VSVG	Dharmacon (V3SH11240-224779127)			
pSMART Lenti-NT shRNA CAG-TurboRFP-VSVG	Dharmacon (VSC11719)			

Supplemental Table 3.6. Source and construction of plasmid vectors.

Chapter 4. Discussion and Future Directions

4.1 Summary of contribution

In this dissertation, we establish that TDP43 homeostasis is critical for cell health and plays important roles in RNA processing and stability. Given the integral role of TDP43 in RNA regulation, Chapter 2 reviews mechanisms that modulate RNA stability and how they are disrupted in neurodegenerative disease. In Chapter 3, we relate two cardinal features of ALS, hyperexcitability and TDP43 pathology, to describe a novel mechanism of TDP43 dyshomeostasis and demonstrate that truncated TDP43 isoforms, sTDP43-1 and -2, recapitulate signature pathologic features of ALS. This work offers potential insight into the late-onset, selective vulnerability of motor neurons in ALS, and reconciles gain- and loss-of-function mechanisms in disease pathogenesis. Further work in Appendix A describes the development of a technique to measure TDP43 homeostasis at the endogenous level via longitudinal fluorescence microscopy, a methodology further detailed in Appendix B. Despite these significant contributions to the ALS field, many questions remain. This chapter outlines outstanding questions and potential strategies to move this work forward.

4.2 Hyperexcitability and alternative splicing

In Chapter 3, we describe activity-dependent *TARDBP* alternative splicing events that result in the inclusion of a small exon encoding a unique 18-amino acid C-terminus. Neuronal

activity is a known regulator of alternative splicing[1], but further studies are required to determine how activity modulates splicing in this context.

The activity-dependent inclusion of small exons is not unique to sTDP43. Long considered to be genetic noise prior to advances in deep genome sequencing and computational methodology[2], recent transcriptome-wide analyses highlight the biological relevance of small exons <20 amino acids in length[3]. Although inclusion or skipping of these microexons represents only 1% of all alternative splicing events, they constitute up to one-third of all conserved neuronally-regulated alternative splicing events between humans and mice[4,5], and 90% of regulated microexons are included at the highest rates in neurons[4]. Moreover, 80-90% of microexons maintain an open reading frame[4,5], allowing their inclusion to alter protein function or localization. Taken together, these studies suggest that microexon inclusion is a conserved neuronal phenomenon that leads to functional changes in proteins throughout the genome[5]. However, activity-dependent regulation of these splicing events is poorly characterized. To date, a small number of RNA-binding proteins including RBFOX, PTBP1, and nSR100 are shown to regulate microexon inclusion[4–6], but future work is required to determine how activity alters these and other regulatory elements relevant to microexon splicing.

Alternative splicing is regulated through a complex system in which the combined effect of both *cis* and *trans* elements determine differential expression of isoforms depending on tissue-type, stage of embryonic development, and other external factors[7–9]. *Cis* elements include enhancers and silencers, or sites where *trans*-acting elements bind to increase or decrease the likelihood that a nearby splice site will be utilized, respectively. Proximity of *cis* elements to the splice site[10,11], their accessibility due to RNA structure[12], and combinatorial effects all determine if splicing occurs. These elements are also highly context dependent, such that an enhancer in one

milieu may function as a silencer in another[7,13]. Intriguingly, some *trans*-acting splicing regulators are, themselves, regulated by neuronal activity. For example, hnRNPA2/B1 is a splicing regulator[14] that is upregulated in response to neuronal activity through impaired unproductive splicing and translation (RUST)[15], a process by which a protein regulates its expression through alternative splicing and destabilization of its own mRNA transcript[16,17]. A striking number of splicing factors and elements of the splicing machinery are regulated through RUST[15,18–23], raising the question of whether hyperexcitability drives global changes in *trans*-acting splicing regulators through alternative splicing (Figure 4.1).

Identification of activity-dependent regulators specific to sTDP43 alternative splicing could be explored in two parallel studies. First, the use of RNA-sequencing (RNA-seq) to identify *trans*-acting splicing factors that are differentially up or downregulated in response to neuronal activity. Numerous studies have employed this strategy to identify activity-modulated genes, in which hyperactivity was induced through chemically induced or electroconvulsive seizures *in vivo*[24–27] or pharmacologically in cultured neurons[28,29]. Though more than 1000 genes are reported to be regulated by neuronal activity in this manner[1,30], we are interested in regulators specific to motor neurons due to their selective vulnerability in ALS and the enrichment of sTDP43-1 and -2 transcripts in this cell type. As such, any motor neuron-specific changes are likely to be diluted by surrounding cell types *in vivo*, and differentially expressed genes in the cultured hippocampal neurons used in most *in vitro* studies may not be applicable to motor neurons. As such, pharmacologically stimulated iPSC-derived motor neuron cultures[31] are the ideal model system, though primary rodent spinal motor neurons are a viable alternative. RNA-seq comparing both differential gene expression and changes in alternative splicing may reveal relevant changes in *trans* regulatory elements, through RUST or some other means. Second, use of predictive

software[32,33] could identify splice-site adjacent *cis*-acting regulatory elements within the *TARDBP* transcript. Moreover, this software could be used to further identify elements that resemble known binding sites of differentially expressed splicing factors identified via RNA-seq. The disease-relevance of any activity-dependent regulatory elements identified through these studies could be further verified through knockdown or upregulation of *trans*-acting elements or steric hindrance of *cis* elements in hyperactive cultures to determine if elevated sTDP43 levels are returned to baseline.

Though there are several avenues of study to identify how neuronal activity regulates *TARDBP* alternative splicing through both *cis* and *trans* elements, future work should also consider the possibility that the alternative splicing events that give rise to sTDP43 are not solely driven by activity. Ongoing studies seek to determine if other forms of stress such as hypoxia, osmotic stress, heat shock, etc. drive sTDP43 formation. If sTDP43 alternative splicing is, indeed, activity dependent, further work should explore sTDP43 pathology in other neurologic disorders that display hyperexcitability such as Alzheimer's Disease[34–36], Parkinson's disease[37,38], and epilepsy[39]. In contrast, if alternative splicing at this site is a generalized response to stress, we could extend this work to include traumatic brain injury, stroke, and other neurological disorders of interest.

4.3 Autoregulation of sTDP43

As described in Chapter 1, TDP43 is regulated through a negative feedback loop in which TDP43 binds the 3' UTR of its own transcript, triggering alternative splicing within the 3' UTR[17,40], mRNA destabilization, and reduced protein expression[16,17,41]. In Chapter 3 we show that sTDP43 is unable to regulate flTDP43 via this mechanism, but it remains unclear if sTDP43 is

regulated by the full-length protein. Although previous studies indicate that only the full-length *TARDBP* transcript is subject to autoregulation[42], work by Polymenidou et al. describes a TDP43 isoform highly resembling sTDP43, termed isoform 3, that is significantly upregulated at both the transcript and protein level by extended TDP43 overexpression[17]. The same study shows a robust increase in isoform 3 following knockdown of the essential nonsense-mediated decay (NMD) component UPF1, suggesting that isoform 3 abundance is regulated by TDP43-mediated alternative splicing and subsequent NMD. Though sTDP43 is not a predicted target of NMD due to loss of the canonical TDP43 stop codon, additional downstream splicing events would render it an NMD substrate.

To determine if flTDP43 regulates sTDP43 in a manner similar to isoform 3, sTDP43 levels could be examined via qRT-PCR, western blotting, and ICC following overexpression of flTDP43. These studies could then be repeated in the absence of UPF1 or other NMD components to determine if the sTDP43 transcript is an NMD target. If inhibition of NMD results in elevated sTDP43, further characterization of downstream splicing events could be accomplished via 3' Rapid Amplification of cDNA ends (RACE), a technique that utilizes the mRNA poly(A) tail as a generic priming site to identify mRNA sequences that result from unknown splicing events within the 3' UTR. RACE performed in the presence and absence of exogenous TDP43 could further verify that flTDP43 mediates sTDP43 splicing, as well as identify any downstream splicing events that render sTDP43 a target of NMD. Finally, disruption of any identified downstream splicing events through mutagenesis of common splice donors or acceptors to determine if they influence sTDP43 levels would conclusively determine whether sTDP43 is an NMD substrate. Together, these studies will provide mechanistic insight into the regulation of sTDP43, and may be extended to determine how this isoform is dysregulated in the context of ALS.

4.4 Exploration of endogeneous sTDP43 functions

sTDP43 isoforms are highly conserved in humans, non-human primates, and lesser mammals at the transcript and protein levels, and this evolutionary conservation suggests that these isoforms could fulfill unknown functions, perhaps involving a compensatory response to chronic neuronal hyperactivity or generalized stress. However, further studies are needed to determine whether native sTDP43 performs essential functions in motor neurons or other cell types.

In Chapter 3, we show that sTDP43 transcripts are significantly enriched in murine and human lumbar motor neurons and are present at low levels in the murine frontal cortex. We also identify sTDP43 pathology in both neurons and astrocytes, indicating that sTDP43 levels vary across cell types. To further characterize cell-type specific sTDP43 transcript levels, examination of previously published RNA-seq data derived from either post-mortem tissue or iPSC-derived cells relevant to ALS (e.g. glia, muscle, von Economo neurons) could reveal the general abundance of sTDP43 in various cell and tissue types. However, given that independent RNA-seq studies are difficult to compare due to differences in sample preparation and sequencing methodology, parallel studies could examine the relative sTDP43 abundance in forebrain-like neurons[43], motor neurons[31], and astrocytes[44] derived from the same line of iPSCs. Identification of relative transcript and protein abundance via qRT-PCR and immunostaining, respectively, may indicate the importance of sTDP43 in each cell type and determine if high sTDP43 transcript levels are specific to motor neurons.

To determine if sTDP43 serves a necessary function in these cell types, CRISPR/Cas9 genomic engineering could be used to disrupt the splice acceptor common to both sTDP43-1 and -2, thereby blocking its formation and creating a sTDP43-null iPSC line. This line could be

differentiated into forebrain-like neurons, motor neurons, or astrocytes, and subsequent survival assays could assess whether the absence of sTDP43 is protective or toxic, shedding light on its importance to cell health in each cell type. It is also possible that the generation of sTDP43 is a compensatory response to stress, and further survival studies performed in the presence of excitotoxic stress or other stressors could determine if sTDP43 formation extends survival. Finally, these iPSC lines could be used to probe non-cell autonomous effects of sTDP43, wherein cocultures of astrocytes and motor neurons with and without the capacity to generate sTDP43 could be assayed for health, survival, and sTDP43 or flTDP43 pathology.

4.5 Further exploration of sTDP43-mediated toxicity

In Chapter 3, we demonstrate that sTDP43 sequesters endogenous TDP43 within cytoplasmic aggregates and induces its clearance from the nucleus, thereby recapitulating signature pathologic changes found in the majority of individuals with ALS. Consistent with these findings, sTDP43 overexpression is toxic to mammalian neurons, suggesting that neurodegeneration results from complementary gain- and loss-of-function mechanisms. However, further studies are required to characterize sTDP43-mediated toxicity.

TDP43 is an essential protein involved in several RNA processing events, and small changes in the localization and expression level of this protein are sufficient to disrupt critical cell processes. As such, cytosolic sequestration and nuclear clearance of TDP43 may contribute to cell death through a loss-of-function mechanism. In addition, sTDP43 inclusions may exert a toxic function of their own through the recruitment and disruption of other proteins and mRNA transcripts. In this manner, sTDP43 may drive neurodegeneration through a gain-of-function mechanism comparable to that proposed for C-terminal TDP43 fragments[45–47]. To assess sTDP43

gain- or loss-of-function toxicity in an unbiased manner, ongoing RNA-seq studies aim to compare the transcriptome of HEK293T cells overexpressing sTDP43 to flTDP43 overexpression or knockdown. Examination of the differentially expressed transcripts and changes in alternative splicing will allow us to determine both the individual transcripts and general pathways in which sTDP43 expression resembles the accumulation or reduction of flTDP43. Further studies are required to determine if restoration of these pathways rescue sTDP43-mediated toxicity. Moreover, given that these studies rely on high levels of exogenous protein expression, the disease relevance of any dysregulated pathways can be determined through comparison to existing RNA-seq data sets from iPSCs, iNeurons, and post-mortem ALS patient tissue.

Though the aforementioned RNA-seq studies are useful in determining ways in which sTDP43 contributes to the loss of TDP43 function, they are limited to the generalized effect of sTDP43 overexpression that is not necessarily specific to cytosolic sequestration of critical cellular components[48]. To further explore the gain-of-function toxicity associated with sTDP43, I suggest two studies to examine a.) proteins, and b.) mRNA transcripts sequestered within these inclusions.

TDP43 reversibly localizes to stress granules in response to various conditions[49–52], and forms reversible homo- and heterotypic interactions with other RNA-binding proteins via its low-complexity domain[49,53,54]. The loss of this low-complexity domain in sTDP43 may alter the dynamics of these protein-protein interactions, leading to irreversible, permissive binding with stress granule proteins and other known TDP43 binding partners[55]. To identify proteins bound in sTDP43 inclusions, we could employ a stable isotope labeling with amino acids in culture (SILAC) strategy in combination with liquid chromatography and tandem mass spectroscopy (LC-MS/MS)[56,57]. To do so, HEK cells overexpressing HaloTagged sTDP43 would be incubated in media containing the “heavy” isotopic forms of arginine and lysine for several days to ensure

complete incorporation of the isotopic amino acids, while HEK293T cells expressing HaloTag alone are incubated in unlabeled "light" media. After sufficient incorporation, both conditions would be mixed in equal parts to reduce experimental variability, and sTDP43 would be immunoprecipitated along with any bound proteins using HaloLink Resin. Samples would then be isolated, digested, and submitted for LC-MS/MS, and any proteins bound to sTDP43-HaloTag at a higher rate than HaloTag alone would be identified based on the mass difference due to the incorporation of the heavy isotopes, indicated by an increase in the mass to charge ratio (m/z) (Figure 4.2A). Mapping these results to known protein sequences[58] would reveal specific proteins bound by sTDP43. Similar studies by Dammer et al. show the enrichment of stress granule components G3BP1, PABPC1, and eIF4A1 in the detergent-insoluble fraction of HEK293T cells overexpressing sTDP43 relative to those overexpressing flTDP43[59], and we verify their finding that exogenous sTDP43 colocalizes with stress granule markers in Chapter 3. Taken together, these data suggest that cytosolic sTDP43 may disrupt normal stress granule dynamics, as sTDP43 lacks the glycine rich domain that mediates reversible protein-protein interactions required for stress granule disassembly[49,53,54]. As such, these proposed experiments would reveal if stress granule components directly interact with cytoplasmic sTDP43 and provide an unbiased means of identifying other components within these inclusions.

In addition to protein, sTDP43 inclusions may bind and sequester mRNA transcripts. TDP43 is an RNA-binding protein and recognizes over one third of all transcribed genes in the human genome[17,40,60,61] through RNA recognition motifs that are intact in sTDP43. Moreover, preliminary results show that cytosolic sTDP43 inclusions colocalize with nucleic acid. To determine which mRNA transcripts are sequestered in these inclusions, HaloTagged sTDP43 or HaloTag alone could be immunoprecipitated from HEK293T cells in conditions that maintain

protein-RNA interactions. Eluted RNA would be submitted for RNA-sequencing to examine the subset of the transcriptome sequestered in these inclusions (Figure 4.2B). If sufficient amounts of RNA cannot be isolated using this protocol, we could employ a candidate-based approach to assess the enrichment of specific transcripts using qRT-PCR. Finally, to assess the contribution of RNA sequestration to neurodegeneration, further work could abolish sTDP43 RNA binding ability to determine if that is sufficient to rescue sTDP43-mediated toxicity[62]. In this manner, these studies would thoroughly characterize the protein and mRNA transcripts sequestered in sTP43 cytosolic inclusions and assess how sTDP43 contributes to neurodegeneration.

4.6 Modulating sTDP43 as a therapeutic strategy for ALS*

Given that we detect sTDP43 in ALS patient tissue and sTDP43 overexpression significantly increases the risk of neurodegeneration in rat cortical neuron models, future work should explore the knockdown or removal of this isoform as a therapeutic strategy for ALS. Antisense oligonucleotides (ASOs), or short, single-stranded oligomers that bind to complementary RNA sequences, can alter protein expression through a variety of mechanisms, including RNase-H recruitment to trigger transcript degradation[58,59] or steric inhibition of translation machinery[60]. However, given that tight regulation of TDP43 levels is essential for cell health, an ideal therapeutic strategy blocks sTDP43 formation and leaves the *TARDBP* transcript intact. Previous work demonstrates that ASOs complementary to splice sites can modulate splicing through the steric inhibition of the spliceosome; ASO-mediated splicing correction was first described as a therapeutic strategy in 1993, and has since been used in a variety of contexts

* This therapeutic strategy was first described in grant proposal by Sami Barmada, and subsequent discussions informed this section.

including diabetes and cancer[61–65]. Recently, and perhaps most notably, ASOs were determined to be remarkably effective in treating children with spinal motor atrophy (SMA) by blocking alternative splicing of the gene *SMN2* to generate sufficient SMN protein and rescue the muscle weakness, limited mobility, and reduced lifespan characteristic of SMA[65–68]. Although the genetic therapeutic targets for SMA are considerably less complicated than those of ALS, we have shown that sTDP43 pathology is driven by hyperexcitability displayed by approximately 80% of ALS patients[69]. If sTDP43 pathology is further verified as a common feature in sporadic and heritable ALS and sTDP43-mediated toxicity is identified as a common mechanism of disease pathogenesis, interruption of its formation could be widely beneficial for a significant number of ALS patients.

To test this therapeutic strategy, we are working in collaboration with Ionis Pharmaceuticals, Inc. to generate ASOs that bind to both splice donors and the shared splice acceptor in *TARDBP* Exon 6 and 3'UTR, respectively. Following identification of ASOs that effectively inhibit sTDP43 formation, their effects on survival in rodent primary neuron or iNeuron models of ALS can be determined via longitudinal fluorescence microscopy. Initial experiments should explore how ASOs alter survival in control neurons to determine if sTDP43 plays a functional role in healthy cells or if it is purely pathogenic. Given that a.) sTDP43 is highly conserved and b.) its transcript is the dominant *TARDBP* species in both control and ALS human spinal motor neurons, its formation is likely not merely an aberrant process. However, its role, if any, in healthy cell function is yet to be determined.

Following establishment of a reliable survival phenotype in iNeurons derived from ALS patients – potentially those with *C9orf72* and *TARDBP* mutations that display initial hyperexcitability[70,71] – we can assess the effect of ASOs on survival via longitudinal fluorescence microscopy. If ASOs effectively rescue toxicity in C9orf72 and TDP43 mutant iNeurons, studies

could be extended to iNeurons derived from sporadic ALS patients or patients with other disease-associated mutations. Such studies could also extend into ALS animal models to determine if ASO treatment extends survival, delays motor deficits, or alters TDP43 pathology.

4.7 Concluding remarks

ALS is a devastating neurodegenerative disorder that affects thousands of people each year. Despite remarkable progress in understanding the genetics and biology that underlie ALS, there are currently no disease-altering therapies available to afflicted patients. In this work, we described a novel pathway of TDP43 dyshomeostasis and developed a technology to better study the metabolism of this protein in a iPSC-derived neuron model of ALS. However, a great deal of work remains, including if and how to exploit this novel pathway to prevent or slow neurodegeneration in ALS. Together with the continued work of myself and others, the findings and proposed experiments outlined in this dissertation are a step toward the identification of safe, accessible, and effective therapies for ALS.

References

1. Hermey, G., Blüthgen, N. & Kuhl, D. Neuronal activity-regulated alternative mRNA splicing. *Int. J. Biochem. Cell Biol.* **91**, 184–193 (2017).
2. Volfovsky, N., Haas, B. J. & Salzberg, S. L. Computational discovery of internal micro-exons. *Genome Res.* **13**, 1216–1221 (2003).
3. Ustianenko, D., Weyn-Vanhentenryck, S. M. & Zhang, C. Microexons: discovery, regulation, and function. *Wiley Interdiscip. Rev. RNA* **8**, (2017).
4. Irimia, M. *et al.* A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523 (2014).
5. Li, Y. I., Sanchez-Pulido, L., Haerty, W. & Ponting, C. P. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res.* **25**, 1–13 (2015).
6. Quesnel-Vallières, M. *et al.* Misregulation of an Activity-Dependent Splicing Network as a Common Mechanism Underlying Autism Spectrum Disorders. *Mol. Cell* **64**, 1023–1034 (2016).
7. Barash, Y. *et al.* Deciphering the splicing code. *Nature* **465**, 53–59 (2010).
8. van der Houven van Oordt, W. *et al.* The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J. Cell Biol.* **149**, 307–316 (2000).
9. Weg-Remers, S., Ponta, H., Herrlich, P. & König, H. Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.* **20**, 4194–4203 (2001).
10. Zhang, X. H.-F., Leslie, C. S. & Chasin, L. A. Dichotomous splicing signals in exon flanks. *Genome Res.* **15**, 768–779 (2005).
11. Lovci, M. T. *et al.* Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* **20**, 1434–1442 (2013).
12. McManus, C. J. & Graveley, B. R. RNA structure and the mechanisms of alternative splicing. *Curr. Opin. Genet. Dev.* **21**, 373–379 (2011).
13. Fu, X.-D. & Ares, M., Jr. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **15**, 689 (2014).
14. Golan-Gerstl, R. *et al.* Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. *Cancer Res.* **71**, 4464–4472 (2011).
15. Kolisnyk, B. *et al.* Cholinergic Regulation of hnRNPA2/B1 Translation by M1 Muscarinic Receptors. *J. Neurosci.* **36**, 6287–6296 (2016).
16. Ayala, Y. M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).
17. 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).
18. Soergel, D. A. W., Lareau, L. F. & Brenner, S. E. *Regulation of Gene Expression by Coupling of Alternative Splicing and NMD*. (Landes Bioscience, 2013).
19. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 189–192 (2003).
20. Green, R. E. *et al.* Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* **19 Suppl 1**, i118–21 (2003).
21. Lareau, L. F., Brooks, A. N., Soergel, D. A. W., Meng, Q. & Brenner, S. E. The coupling of

alternative splicing and nonsense-mediated mRNA decay. *Adv. Exp. Med. Biol.* **623**, 190–211 (2007).
22. Zhou, Y., Liu, S., Liu, G., Oztürk, A. & Hicks, G. G. ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. *PLoS Genet.* **9**, e1003895 (2013).
23. Eom, T. *et al.* NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife* **2**, e00178 (2013).
24. Link, W. *et al.* Somatodendritic expression of an immediate early gene is regulated by synaptic activity. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5734–5738 (1995).
25. Lyford, G. L. *et al.* Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* **14**, 433–445 (1995).
26. Qian, Z., Gilbert, M. E., Colicos, M. A., Kandel, E. R. & Kuhl, D. Tissue-plasminogen activator is induced as an immediate–early gene during seizure, kindling and long-term potentiation. *Nature* **361**, 453–457 (1993).
27. Yamagata, K., Andreasson, K. I., Kaufmann, W. E., Barnes, C. A. & Worley, P. F. Expression of a mitogen-inducible cyclooxygenase in brain neurons: regulation by synaptic activity and glucocorticoids. *Neuron* **11**, 371–386 (1993).
28. Bading, H., Ginty, D. D. & Greenberg, M. E. Regulation of gene expression in hippocampal neurons by distinct calcium signaling pathways. *Science* **260**, 181–186 (1993).
29. Lu, B., Yokoyama, M., Dreyfus, C. F. & Black, I. B. Depolarizing stimuli regulate nerve growth factor gene expression in cultured hippocampal neurons. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6289–6292 (1991).
30. Benito, E. & Barco, A. The neuronal activity-driven transcriptome. *Mol. Neurobiol.* **51**, 1071–1088 (2015).
31. Hester, M. E. *et al.* Rapid and efficient generation of functional motor neurons from human pluripotent stem cells using gene delivered transcription factor codes. *Mol. Ther.* **19**, 1905–1912 (2011).
32. Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. & Krainer, A. R. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* **31**, 3568–3571 (2003).
33. Wang, Z. *et al.* Systematic identification and analysis of exonic splicing silencers. *Cell* **119**, 831–845 (2004).
34. Kazim, S. F. *et al.* Early-Onset Network Hyperexcitability in Presymptomatic Alzheimer's Disease Transgenic Mice Is Suppressed by Passive Immunization with Anti-Human APP/Aβ Antibody and by mGluR5 Blockade. *Frontiers in Aging Neuroscience* **9**, (2017).
35. Minkeviciene, R. *et al.* Amyloid β-Induced Neuronal Hyperexcitability Triggers Progressive Epilepsy. *J. Neurosci.* **29**, 3453–3462 (2009).
36. Ping, Y. *et al.* Linking Aβ42-Induced Hyperexcitability to Neurodegeneration, Learning and Motor Deficits, and a Shorter Lifespan in an Alzheimer's Model. *PLOS Genetics* **11**, e1005025 (2015).
37. Cepeda, C. *et al.* Developmental origins of cortical hyperexcitability in Huntington's disease: Review and new observations. *J. Neurosci. Res.* (2019). doi:10.1002/jnr.24503
38. Cummings, D. M. *et al.* Alterations in cortical excitation and inhibition in genetic mouse models of Huntington's disease. *J. Neurosci.* **29**, 10371–10386 (2009).
39. Bromfield, E. B., Cavazos, J. E. & Sirven, J. I. *Basic Mechanisms Underlying Seizures and Epilepsy*. (American Epilepsy Society, 2006).

40. Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
41. Xu, Y.-F. *et al.* Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J. Neurosci.* **30**, 10851–10859 (2010).
42. D'Alton, S., Altshuler, M. & Lewis, J. Studies of alternative isoforms provide insight into TDP-43 autoregulation and pathogenesis. *RNA* **21**, 1419–1432 (2015).
43. Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78**, 785–798 (2013).
44. Serio, A. *et al.* Astrocyte pathology and the absence of non-cell autonomy in an induced pluripotent stem cell model of TDP-43 proteinopathy. *Proceedings of the National Academy of Sciences* **110**, 4697–4702 (2013).
45. Guo, W. *et al.* An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat. Struct. Mol. Biol.* **18**, 822–830 (2011).
46. Bigio, E. H. *et al.* Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathologica* **125**, 463–465 (2013).
47. Mompeán, M., Baralle, M., Buratti, E. & Laurents, D. V. An Amyloid-Like Pathological Conformation of TDP-43 Is Stabilized by Hypercooperative Hydrogen Bonds. *Front. Mol. Neurosci.* **9**, 125 (2016).
48. Buratti, E. & Baralle, F. E. TDP-43: gumming up neurons through protein–protein and protein–RNA interactions. *Trends Biochem. Sci.* **37**, 237–247 (2012).
49. Colombrita, C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051–1061 (2009).
50. Dewey, C. M. *et al.* TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol. Cell. Biol.* **31**, 1098–1108 (2011).
51. McDonald, K. K. *et al.* TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum. Mol. Genet.* **20**, 1400–1410 (2011).
52. Liu-Yesucevitz, L. *et al.* Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* **5**, e13250 (2010).
53. Gilks, N. *et al.* Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**, 5383–5398 (2004).
54. Guenther, E. L. *et al.* Atomic Structures of Segments from TDP-43 LCD and insight into Reversible and Pathogenic Aggregation. (2018).
55. Freibaum, B. D., Chitta, R. K., High, A. A. & Taylor, J. P. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J. Proteome Res.* **9**, 1104–1120 (2010).
56. Trinkle-Mulcahy, L. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res.* **8**, (2019).
57. Trinkle-Mulcahy, L. *et al.* Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* **183**, 223–239 (2008).
58. Pappin, D. J. C., Hojrup, P. & Bleasby, A. J. Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology* **3**, 327–332 (1993).

59. Dammer, E. B. *et al.* Coaggregation of RNA-binding proteins in a model of TDP-43 proteinopathy with selective RGG motif methylation and a role for RRM1 ubiquitination. *PLoS One* **7**, e38658 (2012).
60. Sephton, C. F. *et al.* Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J. Biol. Chem.* **286**, 1204–1215 (2011).
61. Bhardwaj, A., Myers, M. P., Buratti, E. & Baralle, F. E. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* **41**, 5062–5074 (2013).
62. Flores, B. N. *et al.* An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep.* **27**, 1133–1150.e8 (2019).
63. Kordasiewicz, H. B. *et al.* Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. *Neuron* **74**, 1031–1044 (2012).
64. Smith, R. A. *et al.* Antisense oligonucleotide therapy for neurodegenerative disease. *J. Clin. Invest.* **116**, 2290–2296 (2006).
65. Baker, B. F. *et al.* 2′-O-(2-Methoxy)ethyl-modified Anti-intercellular Adhesion Molecule 1 (ICAM-1) Oligonucleotides Selectively Increase the ICAM-1 mRNA Level and Inhibit Formation of the ICAM-1 Translation Initiation Complex in Human Umbilical Vein Endothelial Cells. *Journal of Biological Chemistry* **272**, 11994–12000 (1997).
66. Dominski, Z. & Kole, R. Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8673–8677 (1993).
67. Bayever, E. *et al.* Systemic Administration of a Phosphorothioate Oligonucleotide with a Sequence Complementary to p53 for Acute Myelogenous Leukemia and Myelodysplastic Syndrome: Initial Results of a Phase I Trial. *Antisense Research and Development* **3**, 383–390 (1993).
68. Marwick, C. First 'Antisense' Drug Will Treat CMV Retinitis. *JAMA* **280**, 871 (1998).
69. Bennett, C. F. & Swayze, E. E. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol.* **50**, 259–293 (2010).
70. Havens, M. A. & Hastings, M. L. Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.* **44**, 6549–6563 (2016).
71. Chiriboga, C. A. *et al.* Results from a phase 1 study of nusinersen (ISIS-SMNRx) in children with spinal muscular atrophy. *Neurology* **86**, 890–897 (2016).
72. Finkel, R. S. *et al.* Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *Lancet* **388**, 3017–3026 (2016).
73. Wood, M. J. A., Talbot, K. & Bowerman, M. Spinal muscular atrophy: antisense oligonucleotide therapy opens the door to an integrated therapeutic landscape. *Hum. Mol. Genet.* **26**, R151–R159 (2017).
74. Talbott, E. O., Malek, A. M. & Lacomis, D. The epidemiology of amyotrophic lateral sclerosis. *Handb. Clin. Neurol.* **138**, 225–238 (2016).
75. Wainger, B. J. *et al.* Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep.* **7**, 1–11 (2014).
76. Devlin, A.-C. *et al.* Human iPSC-derived motoneurons harbouring TARDBP or C9ORF72 ALS mutations are dysfunctional despite maintaining viability. *Nat. Commun.* **6**, 5999 (2015).

Figures

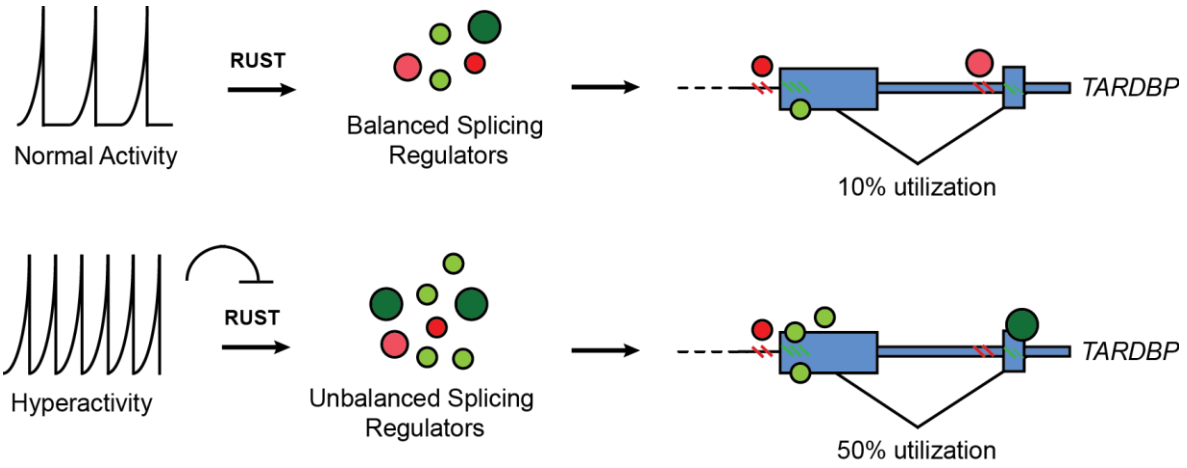


Figure 4.1. Proposed mechanism for activity-dependent regulation of *trans*-acting splicing factors. (A) In healthy cells, regulated unproductive splicing and translation balances levels of various *trans*-acting activators (green) and repressors (red). The abundance and ratio of these factors results in low levels of sTDP43 splicing via binding to enhancers (green lines) and silencers (red lines) within the *TARDBP* transcript. (B) In hyperactive cells, disruption of RUST results in unbalanced activator and repressor homeostasis. As a result, the sTDP43 splice site is more heavily utilized, leading to sTDP43 pathology.

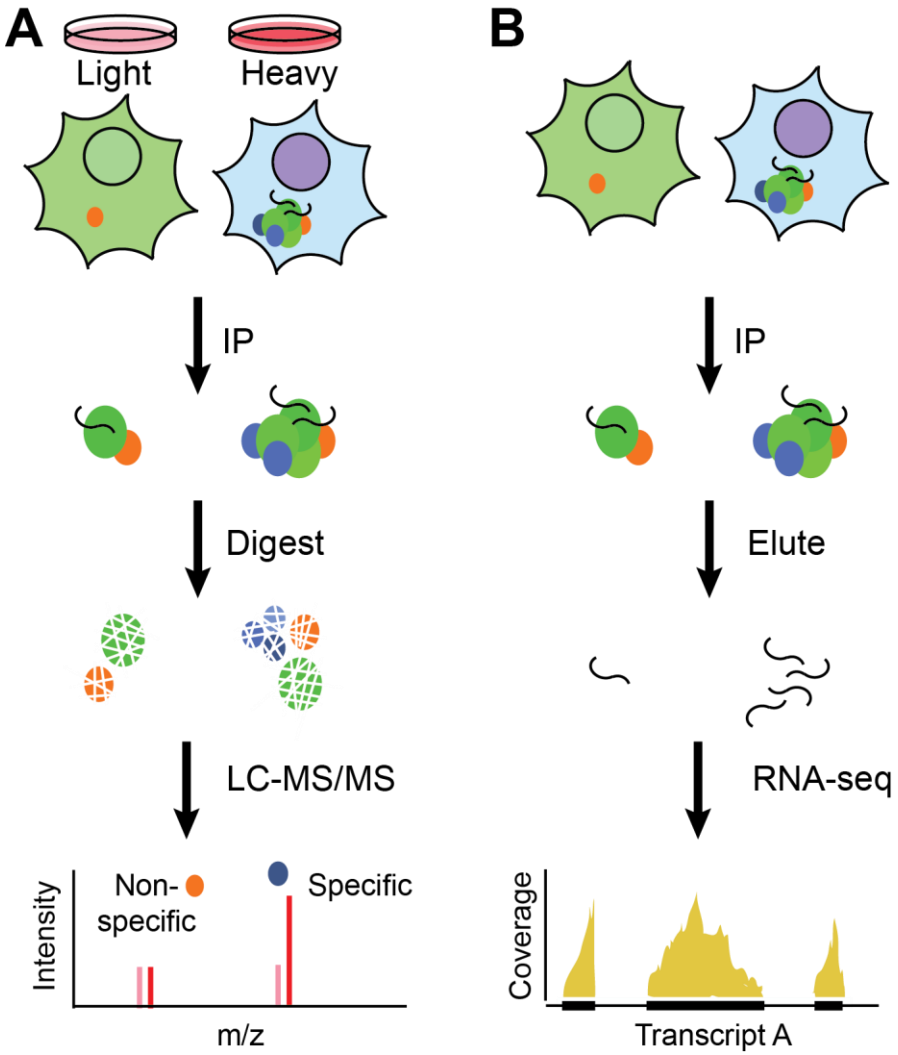


Figure 4.2. Identification of proteins and mRNA transcripts sequestered by sTDP43. (A) To examine sTDP43-protein interactions, HEK293T cells transfected with HaloTagged sTDP43 (green) will be cultured in heavy medium (red), and cells transfected with HaloTag alone will be cultured in light medium (pink). Following immunoprecipitation of the sTDP43-protein complexes (orange, purple), they will be digested with trypsin and submitted for liquid chromatography followed by tandem mass spectroscopy. (B) To identify sTDP43-RNA interactions, HEK293T cells overexpressing HaloTagged sTDP43 or HaloTag alone will be immunoprecipitated. Bound RNA will eluted and submitted for RNA-seq.

Appendix A. Single-cell TDP43 Synthesis and Metabolism in Human iPSC-derived Neurons

A.1 Introduction

Although TDP43 dysregulation is observed in the majority of ALS patients, the mechanisms underlying its accumulation remain unclear. TDP43-positive inclusions in ALS patient tissue are highly ubiquitinated[1,2], suggesting that protein turnover is impaired in disease. In support of this, full-length TDP43 and its truncated 25 kDa and 35 kDa fragments are degraded through both the ubiquitin proteasome system (UPS)[3–5] and autophagy[6–9], and disease-associated mutations confer resistance to TDP43 degradation[10]. Moreover, compounds that enhance TDP43 turnover mitigate toxicity in primary neurons, suggesting that modulating TDP43 metabolism may be a potential therapeutic strategy for ALS[11]. However, most studies exploring TDP43 degradation primarily rely on overexpression models[5,11,12], and given that even slight changes in TDP43 levels are sufficient to disrupt cell function[13], supraphysiologic expression of TDP43 may mask phenomena relevant to human disease.

Here, we established a model system to monitor *endogenous* TDP43 metabolism in neurons differentiated from control and ALS patient-derived induced pluripotent stem cells (iPSCs). Future work seeks to utilize this system to determine if TDP43 turnover is altered in patients harboring mutations in *TARDBP* or *C9orf72*, but this appendix will focus on the development of this technology, characterization of its capabilities, its drawbacks, and how we may apply it in the future.

A.2 Results

A.2.1 Labeling native TDP43 in an iNeuron model of ALS

To investigate TDP43 metabolism in the context of human disease, we established an iPSC-derived neuron (iNeuron) model system in which endogenous TDP43 is tagged with Dendra2, a photoswitchable protein that irreversibly converts from green to red fluorescence following exposure to 405 nm light[14]. To label native TDP43, we utilized CRISPR/Cas9 genome editing to introduce the Dendra2 open reading frame (ORF) into the *TARDBP* locus of control and ALS patient-derived iPSCs (Table A.1). We selected a dual nickase strategy to minimize the risk of unintended insertions and deletions[15], wherein two single-guide RNAs (sgRNAs) directed Cas9(D10A) to induce single-stranded nicks immediately upstream and downstream of the *TARDBP* stop codon. In addition, a vector containing the Dendra2 ORF flanked by 400 bp of homologous sequence 5' and 3' to the *TARDBP* stop codon was introduced to facilitate homology directed repair (HDR), thereby appending Dendra2 to the C-terminus of TDP43 (Figure A.1A). Positive cells were selected based on Dendra2 fluorescence, and unique clones were positively selected and enriched by sequential passaging until a homogeneous population was achieved. PCR and western blotting confirmed the successful insertion of the Dendra2 ORF upstream of the *TARDBP* stop codon, and further indicated that most lines were heterozygous for the insertion (Figure A.1B). Exposing the resultant lines to 405 nm light was sufficient to convert TDP43-Dendra2 fluorescence from green to red, further confirming the presence and functionality of integrated Dendra2 (Figure A.1C).

Next, to ensure reliable and efficient differentiation of iPSCs into iNeurons, we utilized transcription activator-like endonucleases (TALENs) specific for the *CLYBL* safe harbor locus to integrate the transcription factors Neurogenin 1 and 2 (Ngn1-2) under a doxycycline-inducible

promoter (Figure A.1D). Expression of these factors is sufficient to drive the differentiation of iPSCs into iNeurons that display immunocytochemical and electrophysiological properties of glutamatergic, forebrain-like neurons within 2 weeks, as demonstrated by others and described in Chapter 3[16–18] (Figure A.1E,F).

In this manner, we introduced Dendra2 and Ngn1-2 to iPSC lines derived from both healthy controls and ALS patients, such that a.) native TDP43 is fluorescently labeled with the photoconvertible protein Dendra2, and b.) these lines can be reliably and rapidly differentiated into a pure population of forebrain-like neurons.

A.2.2 Tracking TDP43 metabolism and survival at the single-cell level

We next utilized optical pulse labeling (OPL), a technique that enables non-invasive measurements of protein turnover using Dendra2[11,14]. Following a pulse of 405 nm light, we can monitor the loss of photoconverted red signal and the return of newly-synthesized green signal to follow native TDP43 degradation and synthesis, respectively. To aid in these studies, we utilized longitudinal fluorescence microscopy (LFM), a technique that allows us to follow individual cells over time[19]. In this manner, we can determine the rate of TDP43 synthesis and degradation on a single-cell level and prospectively relate TDP43 metabolism to neuronal fate.

For each experiment, we determined baseline TDP43-Dendra2 fluorescence for hundreds of Day 14 iNeurons per condition, photoconverted, and monitored the return of the green signal and the disappearance of the red signal over the following 24 hours (Figure A.2A). Changes in fluorescence level were used to calculate TDP43 half-life and doubling rate for each cell. We observed a high degree of intercellular variability across all experiments, but median TDP43

doubling rate in control iNeurons was determined to be 20.8h (Figure A.2B) and median TDP43 half-life was 32.6h (Figure A.2C), which is comparable to previous findings[20–22].

Following OPL, we utilized LFM to determine the time of death for each cell to relate TDP43 metabolism and survival. We optimized these experiments such that control iNeurons died off gradually over the course of 2-3 weeks (Figure A.2D). Although there is no significant relationship between control iNeuron survival in steady-state TDP43 levels (Figure A.2E) or TDP43 half-life (Figure A.2F), potential relationships may appear when these studies are extended to iNeurons harboring mutations in *C9orf72* or *TARDBP*.

A.3 Discussion

Thus far, we have established a method to monitor native TDP43 metabolism and relate it to cell survival in an iNeuron model system. Doing so required considerable intellectual and emotional endurance, and I highly recommended that anyone who wishes to utilize this system speaks directly to myself or another contributing author to fully understand the challenges summarized below.

The primary challenges were threefold: First, if TDP43 metabolism is altered in ALS, it is predicted to be a subtle effect given that relatively minor changes in TDP43 levels are sufficient to drive neurodegeneration[11,23–25] and the peak age of ALS onset is 50-60 years old[26]. This, in combination with the lengthy and nuanced protocol required for these studies, raises concern that technical variation between replicates may mask disease-specific phenomena. This informed our decision to differentiate iNeurons via the integration of Ngn1-2, which ensures a homogeneous population of iNeurons. We were also rigorous in adhering to a consistent protocol, but unavoidable differences in iPSC passage number, age/strength of the lamp used for

photoconversion, minor environmental changes, etc. may significantly contribute to experimental variability.

Second, LFM requires each population of cells to gradually die within a reasonable time frame. We found that media changes, light exposure, and cell density are all factors that contribute to cell death, and the shearing force of a media change, an extended photoconversion, or low cell density are sufficient to kill iNeurons *en masse*. Conversely, other populations of iNeurons seem to survive indefinitely, or die at a rate too slow to compare toxicity between populations. Ultimately, we identified a combination of substrate, neuron density, photoconversion time, and imaging media that resulted in ~80% cell death within 2 weeks for healthy control iNeurons. However, if these conditions are suitable for iNeurons derived from other iPSC lines remains to be seen.

Third is the challenge of properly visualizing both iNeurons and TDP43-Dendra2. Identifying and tracking individual neurons over several weeks requires culture conditions in which the cells remain adherent and maintain a monolayer. We tested several substrates and identified a combination of polyethyleneimine[27] and laminin that works in most cases. The correct cell density is also critical, as denser plates are far more likely to form clumps, but sparse cultures die at accelerated rates. Moreover, we are experiencing ongoing issues with a signal to noise ratio when monitoring TDP43 metabolism. Endogenous TDP43 tagged with Dendra2 is dim at steady state levels, and only a portion of the fluorescent protein is photoconverted during OPL. However, extended exposure to 405 nm light is toxic, so it is critical to strike a balance between photoconversion efficiency and phototoxicity. We addressed this by developing an imaging media that is both colorless to avoid optical scattering and includes photostable replacements for phototoxic media components. Additionally, we performed all OPL experiments on the

ImageXpress Micro Device that is equipped with a highly sensitive camera. Even so, the low signal to noise ratio complicates our ability to reliably detect subtle changes in native TDP43 synthesis and degradation. Although regenerating multiple iPSC lines would be a considerable undertaking, this issue could potentially be resolved by replacing Dendra2 with a brighter photoconvertible protein, such as Kaede[28] or mEos2[29]. Finally, visualizing native TDP43 is complicated by the presence of autofluorescent material that is often adherent to live cells. Autofluorescence is most significant in the red channel, and this material is often brighter than photoconverted TDP43-Dendra2 itself. Inclusion of any autofluorescent signal can interfere with accurate measurements, and it often overlaps the desired ROI. RFP signal also increases in dying cells, with slight changes in the focal plane, and as cells become more spherical during their normal migration around the plate. These observations further indicate that careful selection of cells and extremely consistent experimental practice are required to minimize artifacts and accurately measure TDP43 metabolism in this model system.

With this in mind, future studies will focus on boosting the signal to noise ratio to increase confidence in our results, as well as the careful characterization of TDP43 synthesis and half-life in iNeurons harboring mutations in *C9orf72* and *TARDBP*.

A.4 Materials and Methods

A.4.1 Generation and maintenance of iPSCs

Fibroblasts were reprogrammed, validated, and maintained as previously described[30,31]. In brief, iPSC lines were cultured in Essential 8 (E8) media (Gibco A1517001) on plates coated with vitronectin (Gibco A14700) diluted 1:100 in Mg^{2+}/Ca^{2+}-free phosphate buffered saline (PBS,

Gibco 14190-144). Cells were passaged every 5d using 0.5 mM EDTA (Sigma E7889) followed by gentle trituration in E8 media with a P1000 pipette. All lines are verified mycoplasma-free on a monthly basis.

A.4.2 Tagging native TDP43 with Dendra2

Endogenous TDP43 was tagged with Dendra2 as described in Chapter 3, where iPSCs transfected with pX335 vectors encoding Cas9(D10A) and sgRNA pairs targeting sequences flanking the *TARDBP* stop codon. Cells were co-transfected with an HDR vector encoding the Dendra2 open reading frame flanked by 400 bp of sequence homologous to that surrounding the *TARDBP* stop codon (in pUC-minus(M), synthesized by Blue Heron, LLC). To generate isogenic TDP43 M337V and A315T mutant lines, the upstream homologous sequence was extended and altered to encode the desired *TARDBP* mutation. Fluorescent cells were selected and successively passaged to generate iPSC colonies in which 100% of cells expressed Dendra2-labeled TDP43.

A.4.3 Integration of Ngn1/Ngn2 cassette into iPSCs

The Ngn1/2 cassette was integrated into iPSC TDP43-Dendra2 lines as described in Chapter 3. In short, iPSCs were transfected in mTESR-1 media (Cell Technologies 85850) with 2.5 µg of donor DNA and 1.25 µg of each targeting construct using Lipofectamine Stem (Invitrogen STEM00003) and screened for red fluorescence. Partially positive colonies were selected and enriched until a homogeneous population was established.

A.4.4 iNeuron differentiation

The differentiation of iPSCs into iNeurons was performed as described in Chapter 3, with the notable exception that cells were plated on Ibidi 96w μ-plates with an optically clear bottom to aid in detecting low TDP43 signal (Ibidi 89626) that were coated with 0.2% polyethyleneimine[27] and 10 μg/ml laminin. In short, iPSCs were split at a low density into 96w plates, and differentiated with the addition of doxycycline, B27 Supplement, N2 Supplement, BDNF, NT3, and laminin over 2 weeks.

A.4.5 Optical pulse labeling

Neurons were transferred from differentiation media into imaging media (1x SOS (M09-50 Cell Guidance Systems), 1x Glutamax Supplement (Gibco 35050-061), 10 ng/ml BDNF, 10 ng/ml NT3, 0.2 μg/ml laminin, and 2 mg/ml doxycycline in Phenol-free Neurobasal-A (Gibco 12349-015)), and imaged with an ImageXpress Micro (Molecular Devices) equipped with a 20x objective lens. Baseline images were taken in brightfield, the GFP channel (Semrock FITC-3540B-NTE-ZERO filter), and the RFP channel (Semrock TxRed-4040C-NTE-ZERO filter) prior to photoconversion to establish background fluorescence levels. Photoconversion was accomplished using a 2s DAPI (Semrock Brightline DAPI-5060-NTE-ZERO filter) exposure, and neurons were immediately reimaged in brightfield, GFP, and RFP. Cells were then imaged at 3h intervals in a recurring loop for 12h, and again at 24h post-conversion, while maintaining 5% CO_2, humidity, and a temperature of 37°C.

A.4.6 Longitudinal fluorescence microscopy and image processing

Neurons were imaged as described previously[32,33] and in Chapter 3, using a Nikon Eclipse Ti inverted microscope with PerfectFocus3a 20X objective lens and either an Andor iXon3 897 EMCCD camera or Andor Zyla4.2 (+) sCMOS camera. A Lambda XL Xenon lamp (Sutter) with 5 mm liquid light guide (Sutter) was used to illuminate samples, and custom scripts written in Beanshell for use in μManager controlled all stage movements, shutters, and filters. Brightfield images were taken at ~24h intervals for 10-20d following OPL until the majority of cells had died or lifted off the plate. Following each experiment, both OPL and LFM images were background subtracted (Rolling Ball radius = 40) and stacked such that each cell could be tracked over time[19]. A custom graphical user interface was developed to manually generate a unique ROI for each neuron to a.) measure changes in green and red fluorescence and b.) identify the time of death for each neuron based on rounding of the soma and degeneration of neuritic processes[19].

A.4.7 Data Analysis

To calculate the TDP43 half-life for each cell, we normalized each timepoint to the cell's pre-conversion fluorescence and subtracted the fluorescent floor using the following equation, where "X_T" represents the RFP signal at a given time point post-conversion.

$$\text{Normalized } X_T = \frac{((\text{Post-conversion} - \text{Pre-conversion}) - (\text{Post-conversion} - X_T))}{(\text{Post-conversion} - \text{Pre-conversion})}$$

The fluorescent intensity was then plotted over time, and slope of the resultant line was used to calculate TDP43 half-life for each individual cell.

A.5 Acknowledgements

This work was supported by National Institutes of Health (NIH), National Institute for Neurological Disorders and Stroke (NINDS) R01-NS097542, National Institute for Aging (NIA) P30 AG053760 (SJB), the University of Michigan Protein Folding Disease Initiative, and Ann Arbor Active Against ALS. We thank Drs. M. Ward and M. Uhler for advice, protocols, and reagents. We also thank the University of Michigan DNA Sequencing Core and the University of Michigan Center for Chemical Genomics. Finally, we thank those that donated the fibroblast that made these studies possible.

A.6 Author Contributions

This work is a highly collaborative, ongoing effort between Kaitlin Weskamp (K.W.), Nathaniel Safren (N.S.), Elizabeth M. Tank (E.M.T.), and Sami J. Barmada (S.J.B.). All parties contributed to conceptualization, formal analysis, and methodology. S.J.B. additionally contributed to supervision, project administration, and funding acquisition. E.M.T. contributed to investigation and technology development. N.S. was responsible for investigation, technology development, and visualization, and K.W. was responsible for investigation, technology development, writing, and visualization. Roberto Miguez (R.M.) was responsible for developing software required for data analysis.

References

1. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
2. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
3. Kim, S. H. *et al.* Potentiation of amyotrophic lateral sclerosis (ALS)-associated TDP-43 aggregation by the proteasome-targeting factor, ubiquilin 1. *J. Biol. Chem.* **284**, 8083–8092 (2009).
4. Urushitani, M., Sato, T., Bamba, H., Hisa, Y. & Tooyama, I. Synergistic effect between proteasome and autophagosome in the clearance of polyubiquitinated TDP-43. *Journal of Neuroscience Research* NA–NA (2009). doi:10.1002/jnr.22243
5. Wang, X. *et al.* Degradation of TDP-43 and its pathogenic form by autophagy and the ubiquitin-proteasome system. *Neurosci. Lett.* **469**, 112–116 (2010).
6. Caccamo, A. *et al.* Rapamycin Rescues TDP-43 Mislocalization and the Associated Low Molecular Mass Neurofilament Instability. *Journal of Biological Chemistry* **284**, 27416–27424 (2009).
7. Filimonenko, M. *et al.* Functional multivesicular bodies are required for autophagic clearance of protein aggregates associated with neurodegenerative disease. *J. Cell Biol.* **179**, 485–500 (2007).
8. Ju, J.-S. *et al.* Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. *J. Cell Biol.* **187**, 875–888 (2009).
9. Wang, I.-F. *et al.* Autophagy activators rescue and alleviate pathogenesis of a mouse model with proteinopathies of the TAR DNA-binding protein 43. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 15024–15029 (2012).
10. Ling, S.-C. *et al.* ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13318–13323 (2010).
11. Barmada, S. J. *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat. Chem. Biol.* **10**, 677–685 (2014).
12. Scotter, E. L. *et al.* Differential roles of the ubiquitin proteasome system and autophagy in the clearance of soluble and aggregated TDP-43 species. *J. Cell Sci.* **127**, 1263–1278 (2014).
13. Weskamp, K. & Barmada, S. J. TDP43 and RNA instability in amyotrophic lateral sclerosis. *Brain Res.* **1693**, 67–74 (2018).
14. Chudakov, D. M., Lukyanov, S. & Lukyanov, K. A. Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2. *Nat. Protoc.* **2**, 2024–2032 (2007).
15. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
16. Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78**, 785–798 (2013).
17. Lam, R. S., Töpfer, F. M., Wood, P. G., Busskamp, V. & Bamberg, E. Functional Maturation of Human Stem Cell-Derived Neurons in Long-Term Cultures. *PLoS One* **12**, e0169506 (2017).

18. Busskamp, V. *et al.* Rapid neurogenesis through transcriptional activation in human stem cells. *Mol. Syst. Biol.* **10**, 760 (2014).
19. Weskamp, K., Safren, N., Miguez, R. & Barmada, S. Monitoring Neuronal Survival via Longitudinal Fluorescence Microscopy. *JoVE (Journal of Visualized Experiments)* e59036 (2019).
20. Pesiridis, G. S., Tripathy, K. & Tanik, S. A 'two-hit' hypothesis for inclusion formation by carboxyl-terminal fragments of TDP-43 protein linked to RNA depletion and impaired microtubule-dependent transport. *Journal of Biological* (2011).
21. Huang, C.-C. *et al.* Metabolism and mis-metabolism of the neuropathological signature protein TDP-43. *J. Cell Sci.* **127**, 3024–3038 (2014).
22. Alber, A. B., Paquet, E. R., Biserni, M., Naef, F. & Suter, D. M. Single Live Cell Monitoring of Protein Turnover Reveals Intercellular Variability and Cell-Cycle Dependence of Degradation Rates. *Mol. Cell* **71**, 1079–1091.e9 (2018).
23. Janssens, J. *et al.* Overexpression of ALS-associated p.M337V human TDP-43 in mice worsens disease features compared to wild-type human TDP-43 mice. *Mol. Neurobiol.* **48**, 22–35 (2013).
24. Wegorzewska, I. & Baloh, R. H. TDP-43-based animal models of neurodegeneration: new insights into ALS pathology and pathophysiology. *Neurodegener. Dis.* **8**, 262–274 (2011).
25. Barmada, S. J. *et al.* Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
26. Kiernan, M. C. *et al.* Amyotrophic lateral sclerosis. *Lancet* **377**, 942–955 (2011).
27. Vancha, A. R. *et al.* Use of polyethyleneimine polymer in cell culture as attachment factor and lipofection enhancer. *BMC Biotechnol.* **4**, 23 (2004).
28. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H. & Miyawaki, A. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12651–12656 (2002).
29. McKinney, S. A., Murphy, C. S., Hazelwood, K. L., Davidson, M. W. & Looger, L. L. A bright and photostable photoconvertible fluorescent protein. *Nature Methods* **6**, 131–133 (2009).
30. Tank, E. M. *et al.* Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat. Commun.* **9**, 2845 (2018).
31. Weskamp, K. *et al.* Neuronal hyperexcitability drives TDP43 pathology by upregulating shortened TDP43 protein isoforms. *bioRxiv* 648477 (2019). doi:10.1101/648477
32. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
33. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *Elife* **7**, (2018).

Figures

Genotype	Line Name	Age at Biopsy	Age of Onset	Gender	Onset	Additional Information
Control	CT-1	58	-	M	-	Healthy
	CT-2	54	-	F	-	Healthy, isogenic control for TDP43 mutant lines
TDP43 ALS	ALS-TDP43(A315T)	54	-	F	-	A315T mutation introduced via CRISPR/Cas9
	ALS-TDP43(M337V)	54	-	F	-	M337V mutation introduced via CRISPR/Cas9
C9 ALS	ALS-C9-1	54	52	M	Lumbar	C9orf72 positive (44 repeats)
	ALS-C9-2	51	49	M	Lumbar	C9orf72 positive (>44 repeats)

Table A.1. Addition information for iPSCs used in these studies.

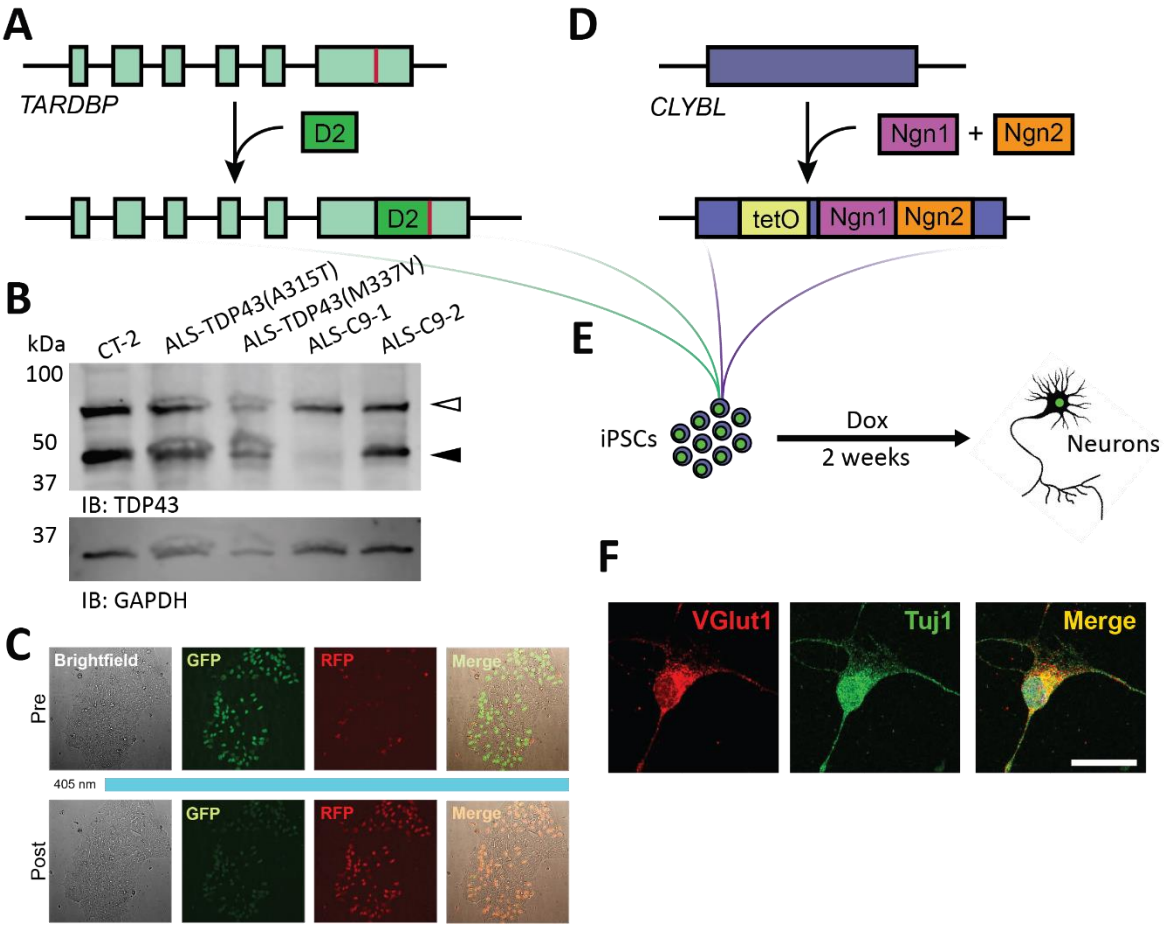


Figure A.1. Establishment of an iNeuron model system to monitor native TDP43 metabolism. (A) Schematic depicting the insertion of Dendra2 upstream of the *TARDBP* stop codon. (B) Immunoblotting reveals that all lines are heterozygous for the insertion of Dendra2, excluding ALS-C9-1, which is homozygous. Black arrow indicates untagged TDP43 at 43 kD, white arrow shows the TDP43-Dendra2 fusion protein at 70 kD. (C) Successful integration of Dendra2 results in the expression of endogenous green TDP43 localized to the nucleus (pre). Following a pulse of 405 nm light (blue box), the fluorescence is shifted from green to red (post). Image depicts a partially-enriched colony. (D) Schematic depicting the insertion of transcription factors Ngn1 and Ngn2 into the *CLYBL* safe harbor locus under a dox-inducible promoter (tetO). (E) Timeline depicting the differentiation of iPSCs into forebrain-like neurons within 2w of doxycycline addition. (F) The resultant neurons express the neuronal markers Vglut1 and Tuj1. Scale bar in (F), 20 μm.

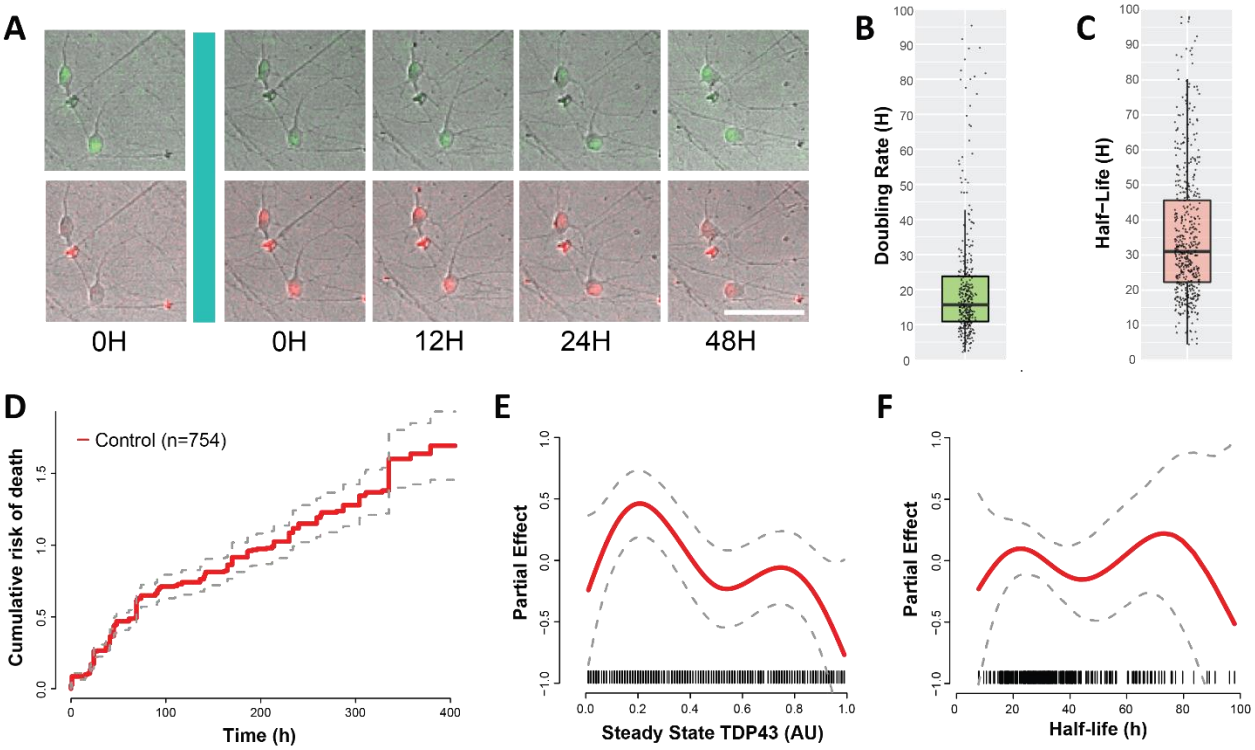


Figure A.2. TDP43 half-life, doubling rate, and survival in control iNeurons. (A) Representative images illustrating optical pulse labeling. At baseline, TDP43-Dendra2 fluoresces green, but following photoconversion (blue line) a portion of the protein is converted to red. Return of the green signal and the degradation of the red signal is evident over time. (B) TDP43 doubling rate in control iNeurons is also highly variable (n=343, median doubling rate=15.61h, S.D. 17.03h). (C) TDP43 half-life in control iNeurons varies widely between individual cells (n=568, median half-life=32.57h, S.D. 33.75h). (D) Representative survival curve depicting the rate of cell death for control iNeurons over the course of several weeks (gray lines are 95% confidence intervals, n=754). Penalized spline analysis did not show a relationship between survival and (E) steady-state TDP43 levels (gray lines are 95% confidence intervals, dashes indicate individual iNeurons, n=307) or (F) TDP43 half-life (gray lines are 95% confidence intervals, dashes indicate individual iNeurons, n=284). Scale bar in (A) 100 μm.

Appendix B. Monitoring Neuronal Survival via Longitudinal Fluorescence Microscopy*

B.1 Abstract

Standard cytotoxicity assays, which require the collection of lysates or fixed cells at multiple time points, have limited sensitivity and capacity to assess factors that influence neuronal fate. The requirement for static "snapshots" of separate populations of cells precludes the ability to discriminate whether an event is a disease driver, a homeostatic response, or merely coincidental. Single-cell longitudinal microscopy overcomes these limitations, allowing the researcher to determine differences in survival between populations and draw causal relationships with enhanced sensitivity. This guide will outline a representative workflow for experiments measuring single-cell survival of rat primary cortical neurons expressing a fluorescent protein marker. The reader will learn how to achieve high-efficiency transfections, collect and process images enabling the prospective tracking of individual cells, and compare the relative survival of neuronal populations using Cox proportional hazards analysis.

* This appendix represents the following manuscript:

Weskamp K, Nathaniel Safren, Roberto Miguez, Sami Barmada. (2018). Monitoring neuronal survival via single-cell longitudinal fluorescence. *Journal of Visual Experiments.* https://doi.org/10.3791/59036

B.2 Introduction

Abnormal cell death is a driving factor in many diseases, including cancer, neurodegeneration, and stroke[1]. Robust and sensitive assays for cell death are essential to the characterization of these disorders, as well as the development of therapeutic strategies for

extending or reducing cellular survival. There are currently dozens of techniques for measuring cell death, either directly or through surrogate markers[2]. For example, cell death can be assessed visually with the help of vital dyes that selectively stain dead or living cells[3], or by monitoring the appearance of specific phospholipids on the plasma membrane[4–6]. Alternatively, measurements of intracellular components or cellular metabolites released into the media upon cellular dissolution are frequently used as proxies for cell death[7,8]. Though these methods provide rapid means of assessing cell survival, they are not without caveats. Each technique observes the culture as a single population, rendering it impossible to distinguish between individual cells and their unique rates of survival. Furthermore, such population-based assays are unable to measure factors that may be important for cell death, including cellular morphology, protein expression, or localization. In many cases, these assays are limited to discrete time points, and do not allow for the continuous observation of cells over time.

In contrast, longitudinal fluorescence microscopy is a highly flexible system that directly and continuously monitors the risk of death on a single-cell basis[9]. In brief, longitudinal fluorescence microscopy involves the transient transfection or transduction of cells with vectors encoding fluorescent proteins. A unique fiduciary is then established, and the position of each transfected cell in relation to this landmark allows the user to image and track individual cells over the course of hours, days, or weeks. When these images are viewed sequentially, cell death is marked by characteristic changes in fluorescence, morphology, and fragmentation of the cell body, enabling the assignment of a time of death for each cell. The calculated rate of death, determined by the hazard function, can then be quantitatively compared between conditions, or related to select cellular characteristics using univariate or multivariate Cox proportional hazards analysis[10]. Together, these approaches enable the accurate and objective discrimination of rates of cell death

among cellular populations, and the identification of variables that significantly predict cell death and/or survival (Figure B.1).

Although this method can be used to monitor survival in any post-mitotic cell type in a variety of plating formats, this protocol will describe conditions for transfecting and imaging rat cortical neurons cultured in a 96-well plate.

B.3 Protocol

B.3.1 Material preparation

1. Dissect cortical neurons from embryonic day 19–20 rat pups and culture at 0.5×10^6 cells per milliliter on poly-D-lysine coated plates for 4 days *in vitro*, as described previously[11–14].
2. Prepare and quantify the plasmid DNA of interest using an endotoxin-free plasmid DNA isolation kit.
3. On *in vitro* day 4 (DIV4), aliquot, filter sterilize, and incubate the following media at 37 °C. Volumes listed are sufficient for transfecting one 96w plate.
    - 6 ml Optimem (OM)
    - 25 ml Neurobasal media (NB)
    - 40 ml NBKY (Neurobasal + 1 mM Kynurenic acid + 10 mM $MgCl_2$, adjusted to a pH of 7.4)
    - 10 ml NBC (Neurobasal + 1x B27 supplement + 1x Glutamax + 1x Pen Strep)

B.3.2 Transfection of rat cortical neurons

1. Modify the provided *Example Transfection Sheet* by adjusting the plate type, plate map, number of DNAs, DNA concentration, and number of wells (green boxes). The total DNA should sum to 0.2 µg per well, regardless of whether one (e.g. DNA A) or multiple (e.g. DNA B and C) DNA constructs are added to each well.
2. Working from the spreadsheet, combine the appropriate amount of OM and DNA in one tube. Combine the appropriate amount of OM and Lipofectamine 2000 in a separate tube.
3. Incubate at room temperature for 5 minutes.
4. Combine the DNA and Lipofectamine OM mixtures and incubate at room temperature for 20 minutes.
5. During this incubation step, use a multichannel pipette and sterile plastic troughs to wash cells 2x with 100 µl per well of NB. Reserve the conditioned media (CM) and store at 37 °C. For this and following steps, take care to minimize the amount of time neurons are exposed to air.
6. Remove the NB media and replace with 100 µl per well of NBKY.
7. After 20 minutes have passed, add 50 µl of the lipofectamine/DNA mixture dropwise to each well.
8. Incubate cells with the Lipofectamine/DNA complexes for 20 min at 37 °C.
9. Rinse 2x with NBKY, and replace with 100 µl of CM and 100 ul of NBC per well.
10. Incubate the plate at 37 °C overnight, and use a fluorescent microscope to check the transfection the next morning. This technique results in an overall transfection efficiency of 5 to 10%.

B.3.3 Imaging

1. Place the plate on a fluorescent microscope with a motorized stage, and establish a fiduciary (e.g. a mark on the bottom of the plate) that will allow you to align the plate each time it is imaged. Save an image of this fiduciary for reference.
2. Navigate to a field of interest, and note the x-y coordinates relative to the fiduciary.
3. Focus on fluorescently labeled cells.
4. Take fluorescent images in the appropriate channel or channels, either manually or in an automated manner. By taking several images at regularly-spaced intervals, a montage of the well can be assembled during image processing (see below).
5. Repeat this process as often as required, aligning to the original fiduciary each time. For survival analysis, imaging takes place every 6-24 h, depending on the cell type and the purpose of the experiment.

B.3.4 Image Processing

Following image acquisition, a series of processing steps are required prior to image analysis. These include, but are not limited to, stitching, stacking, and background subtraction (Figure B.1). The goal of these steps is to produce an image stack, or time series, in which cells are clearly discernible from their background and easy to follow over multiple time points. A brief description of each step, together with a dedicated FIJI macro (*Image_Processing*), is provided below. Annotation within this macro indicates which parameters the user should modify to match the particular needs of their analysis, all of which are contained within lines 1-14 of the code. Additionally, the raw data or input directory should be formatted as shown in Figure B.2. Once

started, *Image_Processing* will automatically advance through stitching, stacking, and background subtraction.

B.3.4.1 Stitching

If a montage of images is taken, stitching can be performed to create a single, larger image for each field of view. For most applications it is preferable to perform stitching prior to stacking. If only one image is taken per well, there is no need to perform this step. Before using the *Image_Processing* macro, it is essential to determine the order in which the images were acquired. To test this, manually stitch a montage of images in FIJI using *Plugins* → *Stitching* → *Grid/Collection stitching*. Adjust the dropdown menus "Type" and "Order" until an accurately stitched image is produced. Adjust the "GRID_TYPE" and "STITCH_ORDER" variables in lines 8 and 9 of the *Image_Processing* macro to match these selections.

B.3.4.2 Stacking

Rather than tracking cells over time across separate image files, stacking can be performed to align consecutive images into a single time series, analogous to a stop frame animation. With successful fiduciary alignment, the individual frames comprising the stacked image should match up well. However, if there are noticeable shifts or rotations between frames, image registration is needed. The *Image_Processing* macro automatically performs registration using the FIJI plugin "MultiStackReg." This should help reduce small misalignments between imaging runs. However, with significant shifts, manually cropping and realigning images may be required. For stacking to run properly, it is critical that the listed time points are contiguous (i.e T1, T2, T3). If the time points are not contiguous (i.e T1, T3) the macro will crash.

B.3.4.3 Background Subtraction (optional)

One potential issue that may arise during image acquisition is uneven illumination. This will result in variations in signal intensity across an image that can confound estimates of fluorescence intensity. In these instances, intensity variations can be eliminated by background subtraction techniques. These are particularly relevant with low signal to noise ratios, where intensity shifts due to uneven illumination can be comparable in magnitude to the signal of the fluorophore itself. There are many background subtraction algorithms, several of which have associated FIJI plugins. The choice of which algorithm to use depends on the properties of the image itself and the signal being measured. Within the FIJI macro *Image_Processing,* the user is given the option to perform "rolling ball" background subtraction on a stacked set of images (line 14). In this method, a local background is determined for every pixel based on the average intensity of a circle surrounding that pixel. This value is then subtracted from the pixel's initial value. The radius of the circle used for local background estimates should be set to at least the diameter of the largest foreground object in the image. This radius can be adjusted in line 15 of the *Image_Processing* macro.

B.3.5 Scoring cell death

Following the processing steps outlined above should yield images that enable the assessment of single-cell survival. The criteria for determining cell death are crucial, and may vary depending on cell type. Here, we outline a set of criteria that we have established for scoring cell death in rat primary cortical neurons, and demonstrate how these criteria can be applied to measure

the rate of cell death for a given population. Three main criteria are used in the identification of dead neurons (Figure B.3):

1. loss of fluorescence intensity (e.g. Neuron 1 at 69 h)
2. rounding of the cell body (e.g. Neuron 2 at 188 h)
3. loss of neurite integrity or blebbing (e.g. Neuron 2 at 188 h)

For accurate comparisons between populations, it is essential that these criteria be applied consistently across the entire dataset. Furthermore, individuals scoring cell death should be blinded to the experimental groups under investigation to eliminate potential sources of bias. Depending on the specific criteria and their generalizability, they may be incorporated into automated algorithms for the unbiased assessment of cellular survival[13–17]. To avoid counting a cell twice, use the "point tool" within FIJI to individually label each cell with a number. Pressing "t" after each point will add the cell identifier to the ROI (region-of-interest) Manager. The identifiers can be visualized by clicking the "labels" and "show all" checkboxes in the ROI Manager.

In the context of survival analysis and other time-to-event analyses, there are three possible outcomes:

1. The event (cell death) has occurred, and the time at which the event occurred is recorded.
2. The event did not occur during the time frame of observation. These observations are censored at the completion of the experiment.

3. The event could not be scored because the cell moved out of the field of view, or was obscured by nearby cells. In this case, the cell is censored when it can no longer be accurately tracked.

For outcome #1, the precise timing of cell death may be difficult to determine based on the imaging interval. For instance, a cell that is alive initially but marked as dead 24 h later may have died at any point within that 24-hour period. To be conservative, we recommend recording the time of death as the last time a cell can be confidently identified as alive (left censoring).

For rapid, robust and consistent determinations of cellular survival and death rates, we employ the "survival" package in R, an open source statistics program (https://www.rstudio.com). Therefore, when recording survival data, it is important to do so in a way that is compatible with subsequent survival analysis in R. A representative spreadsheet (*Survival_spreadsheet.csv*) is included that can be used as a template. In this spreadsheet each cell occupies a single row. The unique identifier for each cell (ID) consists of its corresponding well and ROI number within that well. tp_death is the last time point a cell is observed to be alive, while time_death represents the actual time of death in hours. The time of death will necessarily vary depending on the interval between imaging time points. Finally, the censored status of each cell is recorded in the last column. Here, due to the peculiar way censoring is handled by R, censored cells are marked by "0", while uncensored cells are marked by "1". Note that all cells that live to the last time point are censored, and therefore marked as "0".

B.3.6 Performing Cox proportional hazards analysis and visualizing results

We have included an R script called *survival.R* that can be used to analyze survival data compiled from the above steps. This script allows you to compare the risk of death among populations and their statistical significance using Cox proportional hazards analysis (Figure B.4A), and also plot results as either a Kaplan Meier curve (Figure B.4B) or a cumulative risk of death plot (Figure B.4C). For more information on survival analysis, Cox proportional hazards analysis, and the "survival" package in R, we refer the reader to the 1987 paper by Erik Christensen[10] and https://cran.r-project.org/web/packages/survival/survival.pdf. We will instead focus on the interpretation of the Cox summary table and accompanying plots.

The table generated upon running lines 7 and 8 of the survival.R code provides a summary of the Cox proportional hazards analysis. Four particularly important statistics in the table are highlighted in Figure B.4A. The number in **Box 1** represents the hazard ratio for the group "Mutant" relative to "WT". You may notice that the "WT" group is not listed. This is because the "WT" group serves as the reference population— the risk of death observed in all other groups are compared to that of the reference population to calculate the hazard ratio. Therefore, hazard ratios greater than 1 indicate a faster rate of death in comparison to the reference population, and values less than 1 represent a reduced rate of death. In the example provided, mutant cells display a hazard ratio of 2.2, meaning that they died 2.2x faster than WT cells. By default, R will arrange the groups in alphanumeric order, with the top group serving as the reference population. Placing numbers in front of your group names is an easy way to establish the order in which they are evaluated. The values in **Boxes 2** and **3** represent the p-values and 95% confidence intervals for the hazard ratios, respectively, calculated by Cox proportional hazards analysis. In **Box 4**, the results of the log-rank test are reported. This test evaluates whether there is a statistically significant difference in survival

among populations being tested, but does not describe which groups are different from the others, and does not calculate a magnitude for the observed difference.

Kaplan Meier curves (Figure B.4B) are widely used in clinical trials for evaluating the effects of an intervention on patient survival. For this reason, many researchers are familiar with interpreting survival data visualized this way. In the context of single-cell survival, these plots depict the fraction of cells alive over time in each group. Rather than plotting cell survival, an alternative approach is to depict the rate of cell death in each group via a cumulative risk of death plot (Figure B.4C). In most survival studies, the number of events does not follow a linear progression; rather, for a given rate of death a greater number of events is observed at earlier times. For example, in a population of 100 cells, if 20% of cells die between intervals, then 20 cells will die within the first interval, 16 during the second interval, 13 during the third interval, and so on. This logarithmic trend is conceptually easier to visualize using cumulative risk of death plots, since the y-axis represents the negative log transform of cellular survival. Alternatively, the y-axis of the cumulative risk of death plot can also be presented as % cell death, calculated as $1-1/e^{\text{cumulative risk of death}}$. These plots also enable straightforward comparisons of the risk of death between populations. The magnitude of the hazard ratio reflects the slope of the cumulative risk of death plot for each population, relative to that of the reference group.

B.4 Representative Results

Using the transfection procedure described here, DIV4 rat cortical neurons were transfected with a plasmid encoding the fluorescent protein mApple. 24 h post-transfection, cells were imaged by fluorescence microscopy every 24 h for 10 consecutive days. The resultant images were organized as indicated in Figure B.2, then stitched, stacked, and scored for cell death (Figure

B.1). Figure B.3 shows a time course for 3 representative neurons, two of which die during the course of the experiment (Neurons 1 and 2) while the third survives (Neuron 3). Survival data were analyzed using the R script provided (*survival.R*), and the results summarized in Figure B.4. As depicted by Kaplan Meier (Figure B.4B) and cumulative risk of death plots (Figure B.4C), mutant neurons exhibited a significantly higher risk of death than did WT cells.

B.5 Discussion

Here, we demonstrate a method to directly monitor neuronal survival on a single-cell basis. In contrast to traditional assays for cell death that are limited to discrete time points and entire populations of cells, this method allows for the continuous assessment of a variety of factors such as cellular morphology, protein expression, or localization, and can determine how each factor influences cellular survival in a prospective manner.

This highly-flexible system can be modified to fit a wide array of experimental needs. The frequency and duration of imaging can be easily adjusted, and any protein of interest can be co-transfected with the fluorescent marker to model disease states or investigate protein function[11-15]. Though this article describes the optimal procedure for transfecting rat cortical neurons, the experimental schema could be applied to any post-mitotic cell type. However, the optimal transfection conditions may need to be optimized on a per-cell line basis, and substrates may need to be adjusted to prevent cells from clumping or moving too much to reliably track.

This assay can also be used to relate a variety of neuronal features to survival. Generation of an ROI around the cell body and/or nucleus enables the user to longitudinally monitor cell size and morphology, protein expression level and localization, or the formation of subcellular structures such as puncta or protein aggregates[11–17]. Importantly, because each of these factors is

observed in relation to cell death, it is possible to quantitatively determine how well individual factors predict cellular survival or death during the given time frame. Protein metabolism and cellular pathways may also be assayed by expressing fluorescent reporters that provide real-time measurements of underlying cellular physiology (e.g. gCaMP6f to assay activity). By employing this powerful approach, factors that drive cellular maintenance, function, and dysfunction can be uncovered and studied in detail, thereby inspiring new avenues of inquiry.

B.6 Acknowledgements

We thank Steve Finkbeiner and members of the Finkbeiner lab for pioneering robotic microscopy. We also thank Dan Peisach for building the initial software required for image processing and automated survival analysis. This work was funded by the National Institute for Neurological Disorders and Stroke (NINDS) R01-NS097542, the University of Michigan Protein Folding Disease Initiative, and Ann Arbor Active Against ALS.

B.7 Author Contributions

K.W. and N.S. were equally responsible for conceptualization, methodology, investigation, formal analysis, writing, and visualization. S.J.B. contributed to conceptualization, methodology, formal analysis, writing, visualization, supervision, project administration, and funding acquisition. R.M. was responsible for software.

References

1. Lockshin, R. A. & Zakeri, Z. Cell death in health and disease. *J. Cell. Mol. Med.* **11**, 1214–1224 (2007).
2. Kepp, O., Galluzzi, L., Lipinski, M., Yuan, J. & Kroemer, G. Cell death assays for drug discovery. *Nat. Rev. Drug Discov.* **10**, 221–237 (2011).
3. Lemasters, J. J. *et al.* The mitochondrial permeability transition in cell death: a common mechanism in necrosis, apoptosis and autophagy. *Biochim. Biophys. Acta* **1366**, 177–196 (1998).
4. Vermes, I., Haanen, C., Steffens-Nakken, H. & Reutelingsperger, C. A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. *J. Immunol. Methods* **184**, 39–51 (1995).
5. Chien, K. R., Abrams, J., Serroni, A., Martin, J. T. & Farber, J. L. Accelerated phospholipid degradation and associated membrane dysfunction in irreversible, ischemic liver cell injury. *J. Biol. Chem.* **253**, 4809–4817 (1978).
6. Zwaal, R. F. A., Comfurius, P. & Bevers, E. M. Surface exposure of phosphatidylserine in pathological cells. *Cell. Mol. Life Sci.* **62**, 971–988 (2005).
7. Mitchell, D. B., Santone, K. S. & Acosta, D. Evaluation of cytotoxicity in cultured cells by enzyme leakage. *J. Tissue Cult. Methods* **6**, 113–116 (1980).
8. Moran, J. H. & Schnellmann, R. G. A rapid beta-NADH-linked fluorescence assay for lactate dehydrogenase in cellular death. *J. Pharmacol. Toxicol. Methods* **36**, 41–44 (1996).
9. Arrasate, M. & Finkbeiner, S. Automated microscope system for determining factors that predict neuronal fate. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3840–3845 (2005).
10. Christensen, E. Multivariate survival analysis using Cox's regression model. *Hepatology* **7**, 1346–1358 (1987).
11. Barmada, S. J. *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat. Chem. Biol.* **10**, 677–685 (2014).
12. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
13. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *Elife* **7**, (2018).
14. Archbold, H. C. *et al.* TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci. Rep.* **8**, 4606 (2018).
15. Green, K. M. *et al.* RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat. Commun.* **8**, 2005 (2017).
16. Park, S.-K. *et al.* Overexpression of the essential Sis1 chaperone reduces TDP-43 effects on toxicity and proteolysis. *PloS Genet.* **13**, e1006805 (2017).
17. Gupta, R. *et al.* The Proline/Arginine Dipeptide from Hexanucleotide Repeat Expanded C9ORF72 Inhibits the Proteasome. *eNeuro* **4**, (2017).

Figures

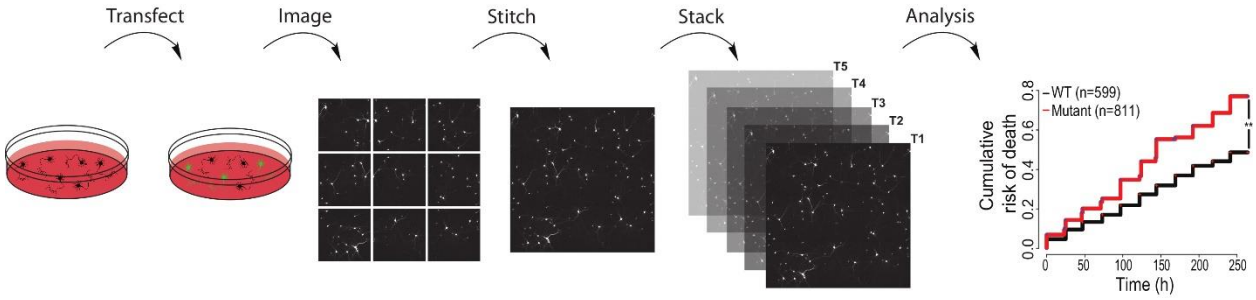


Figure B.1. Schema for a typical survival experiment. Rat cortical neurons are transfected at DIV4 using the procedure outlined in this article. Beginning 24 h post-transfection, cells are imaged at regularly spaced intervals in accordance with the specific requirements of the experiment. Images are stitched and stacked before cell death is scored, and Cox proportional hazard analysis is used to compare the risk of death between populations.

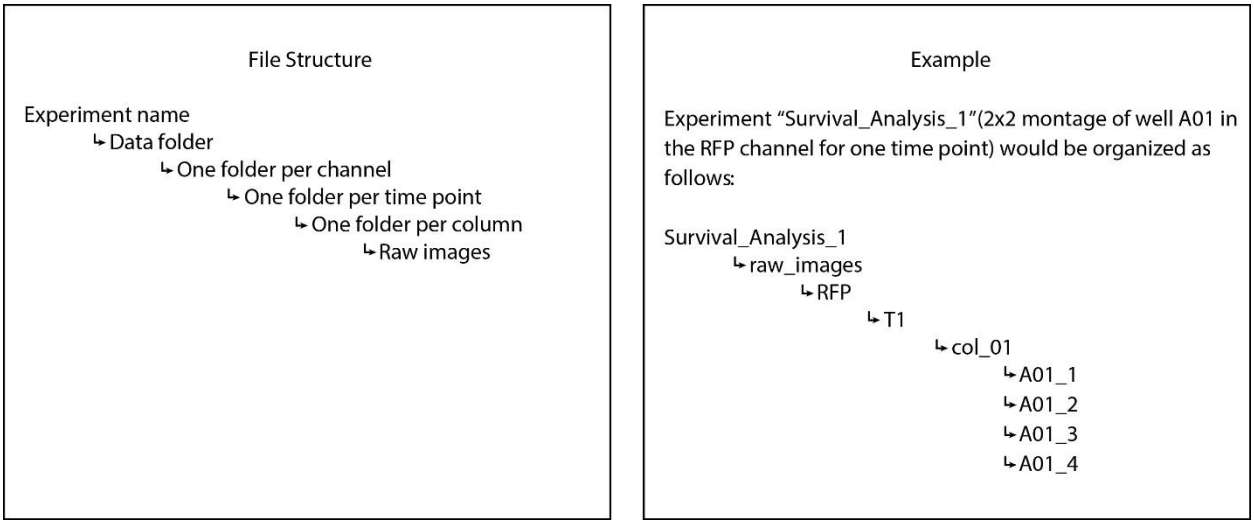


Figure B.2. Required file structure. The provided FIJI macro requires that the raw data are formatted in a specific way. To utilize *Image_Processing*, organize the raw data as shown on the left. An example experiment and accompanying file structure is shown on the right.

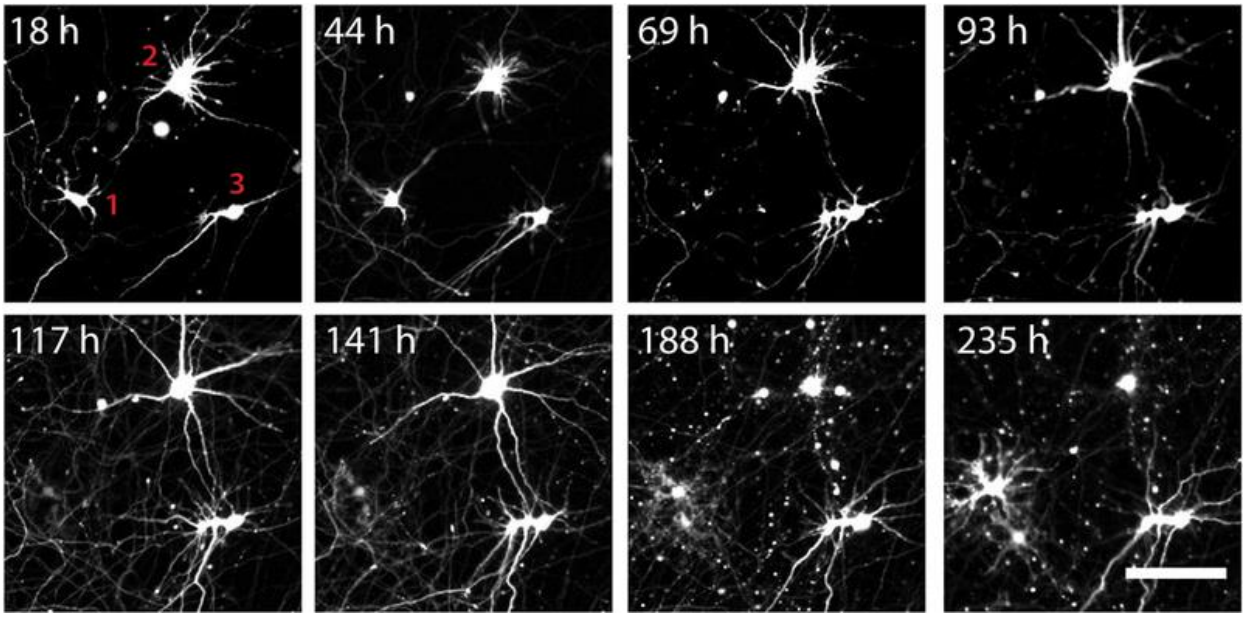


Figure B.3. Scoring cell death in transfected rat cortical neurons. Using the methods described in this article, rat cortical neurons were transfected with a plasmid encoding the fluorescent protein mApple. Cells were then imaged approximately every 24 h, the images were stitched and stacked, and cell death scored using the criteria provided. Cell death is indicated for Neuron 1 at 69 h, as evidenced by loss of fluorescence. Neuron 2 dies at 188 h, as indicated by fragmentation of the processes and rounding of the cell body. Neuron 3 survives for the duration of the experiment. Scale bar = 50 μm.

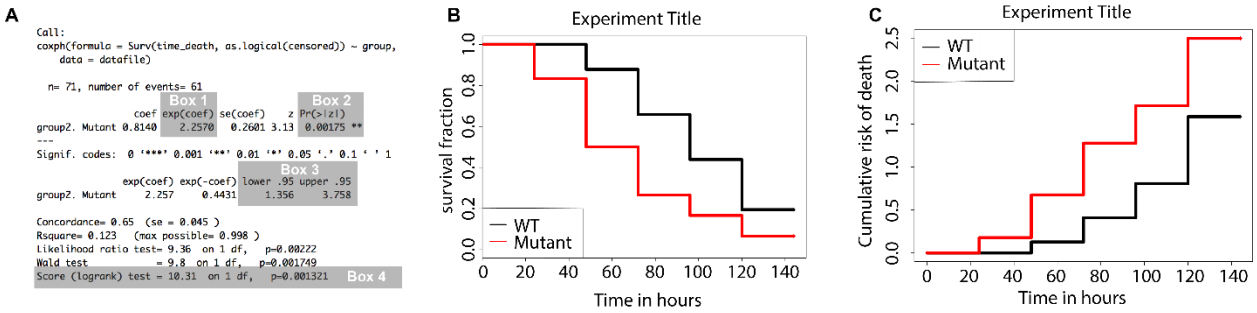


Figure B.4. Interpretation of Cox proportional hazard analysis. (A) The output summary includes four important statistics that are highlighted in this figure. Box 1 includes the hazard ratio of the experimental group relative to the control group, while box Box 2 and 3 show the p-values and 95% confidence interval for each hazard ratio, respectively. Box 4 highlights the results of the log-rank test. These data are also depicted via a Kaplan Meier curve (B) and a cumulative risk of death plot (C).

Appendix C. Establishing a Medium-Throughput Screen in Collaboration with Verge Genomics

C.1 Introduction

Despite years of effort, therapeutic options for ALS remain limited[1]. Conventional methods of drug development using animal models and immortalized cell systems are widely available, but differences in species, cell type, and genetic context complicate the identification of therapeutics relevant to human disease[2]. Though the use of post-mortem patient tissue can avoid these shortcomings, it is often difficult to obtain and reflects only the end-stage of disease. Hence, human induced pluripotent stem cell (iPSC) models have emerged as an alternative method for drug screens. These cells reflect the complicated genetic signature of human patients, including those in which the underlying genetic cause is unknown. Moreover, their supply is not limited given that iPSCs are dividing cell lines. Finally, iPSCs can be differentiated into multiple relevant cell types, allowing for a model system tailored to the subset of cells affected in disease.

As iPSC technology has advanced, several research groups and biotech companies have explored the use of these cells to screen for novel therapeutics. Among them is Verge Genomics, a startup that utilized machine learning and transcriptomics data from C9ORF72 and sporadic ALS post-mortem patient tissue, SOD1 mice, and *in vitro* cell culture data to identify a network of 295 genes down-regulated in diseased samples. They then identified a library of compounds to modulate these novel targets. To screen the efficacy of these compounds, we established a

medium-throughput screen in an iPSC-derived motor neuron (iMN) model of ALS for use in tandem with our longitudinal fluorescence microscopy platform. Although it would be interesting to comment on the mechanisms of these compounds, their identities remained strictly proprietary. As such, Appendix C will focus on the establishment of this screen and its potential future applications.

C.2 Establishing a longitudinal fluorescence microscopy screen in iPSC-derived motor neurons

C.2.1 Utilization and modification of BrainXell motor neuron precursors

We purchased cells from BrainXell, a company that generates spinal motor neuron precursors from patient-derived iPSCs using a proprietary, directed small molecule differentiation system[3,4]. The precursors are cryopreserved at day 28 *in vitro*, shipped to the user, and can fully differentiate into motor neurons expressing Hb9, Isl1, FoxP1, and ChAT within one week of plating[3,4]. We utilized several lines derived from either control or SOD1-ALS patients (Table C.1). Although these cells differentiated rapidly and stained positive for both motor neuron and panneuronal markers, they formed clumps in culture that made it difficult to track their survival over time via longitudinal fluorescence microscopy (Appendix B). However, we eventually determined that the omission of Geltrex, a product that mimics the basement membrane matrix, aids in the maintenance of a cellular monolayer. We also modified the differentiation timeline to accommodate viral transduction, which allowed us to fluorescently label and track individual cells throughout the experiment (Figure C.1).

C.2.2 Transduction with a fluorescent marker

For these studies, we utilized a commercially available lentivirus in which EGFP is driven by an eF1α promoter (VectorBuilder, VB170224-1047xty). Initial experiments revealed that this virus transduced very efficiently (Figure C.2A), and a multiplicity of infection (MOI) of 1 was determined to be sufficient for future studies. The virus was applied to the cells at Day 2 for 12-16 hours, and expression was evident within 24 hours (Figure C.1).

C.2.3 Longitudinal fluorescence microscopy

One week after plating (Day 7), the EGFP-positive neurons were imaged once every ~24 hours for 10 days as described in Appendix B. Fortunately, and atypically for iNeurons, these cells move very little and were easily tracked with our in-house, automated software to determine their time of death in an unbiased and high-throughput manner[5–9].

C.2.4 Identification of a phenotype

Under baseline conditions, in which all lines were maintained in conditioned growth media, both the control and disease lines died very slowly, and at remarkably similar rates (Figure C.2B). We next sought to identify conditions that would reveal a disease-specific phenotype, including the application of glutamate to induce excitotoxicity and the gradual withdrawal of nutrients. While glutamate alone had no effect on survival (Figure C.2B), we found that glutamate in combination with complete replacement of the media with Hanks' Balanced Salt Solution (HBSS) resulted in the neurons dying gradually over the course of 10 days, as well as a distinguishable difference between control and SOD1-ALS cell survival (Figure C.2C). Further studies showed that complete media withdrawal into HBSS alone was sufficient to induce reliable cell death, and

examination of individual lines revealed that the SOD1-ALS line ND50010 exhibited the most robust phenotype (Figure C.2D). As such, ND50010 and the control line ND35660 were selected for future studies.

C.2.5 Drug application

Drugs were typically applied at Day 7 and left on for the duration of the experiment. However, the time of drug application and whether or not it was removed varied between experiments. Multiple doses of each drug were diluted in HBSS and added directly to the cells via a full media change. All studies included a DMSO vehicle control, ideally volume matched to the majority of the drugs and never exceeding 0.2% of the well volume. Later, the pan-caspase inhibitor Q-Vd-Oph (2 μM) was identified as a positive control and included in future studies (Figure C.3A).

C.3 Results

While the identity of these drugs remains proprietary, we tested over a hundred different compounds identified by Verge Genomics as potential therapies for ALS. Though most proved to be toxic or ineffective at rescuing disease-associated toxicity, VRG106 showed a modest reduction in hazard ratio, though it was not consistent between experiments (Figure C.3B). For future work, modifications to this assay or different drugs and targets are required to identify novel therapeutics for the treatment of ALS.

References

1. Petrov, D., Mansfield, C., Moussy, A. & Hermine, O. ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment? *Front. Aging Neurosci.* **9**, 68 (2017).
2. Allen, D. D. *et al.* Cell lines as in vitro models for drug screening and toxicity studies. *Drug Dev. Ind. Pharm.* **31**, 757–768 (2005).
3. Li, X.-J. *et al.* Directed differentiation of ventral spinal progenitors and motor neurons from human embryonic stem cells by small molecules. *Stem Cells* **26**, 886–893 (2008).
4. Du, Z.-W. *et al.* Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nat. Commun.* **6**, 6626 (2015).
5. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7821–7826 (2015).
6. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *Elife* **7**, (2018).
7. Barmada, S. J. *et al.* Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
8. Archbold, H. C. *et al.* TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci. Rep.* **8**, 4606 (2018).
9. Flores, B. N. *et al.* An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep.* **27**, 1133–1150.e8 (2019).

Figures

Cell Line Name	Sex	Age	Genotype	Diagnosis
CS809CTR-Tn1	F	48	Control	N/A
CS201iCTR-NTn4	F	56	Control	N/A
ND35660	F	50	Isogenic Control, SOD1 D90D	N/A
ND39030	F	50	SOD1 D90A	ALS, lower-limb
ND35671	F	65	SOD1 A4V	ALS, lower-limb
ND50010	F	44	SOD1 A4V	Asymptomatic

Table C.1. Details on the gender, age, and genotype of patient-derived iPSCs utilized in these studies.

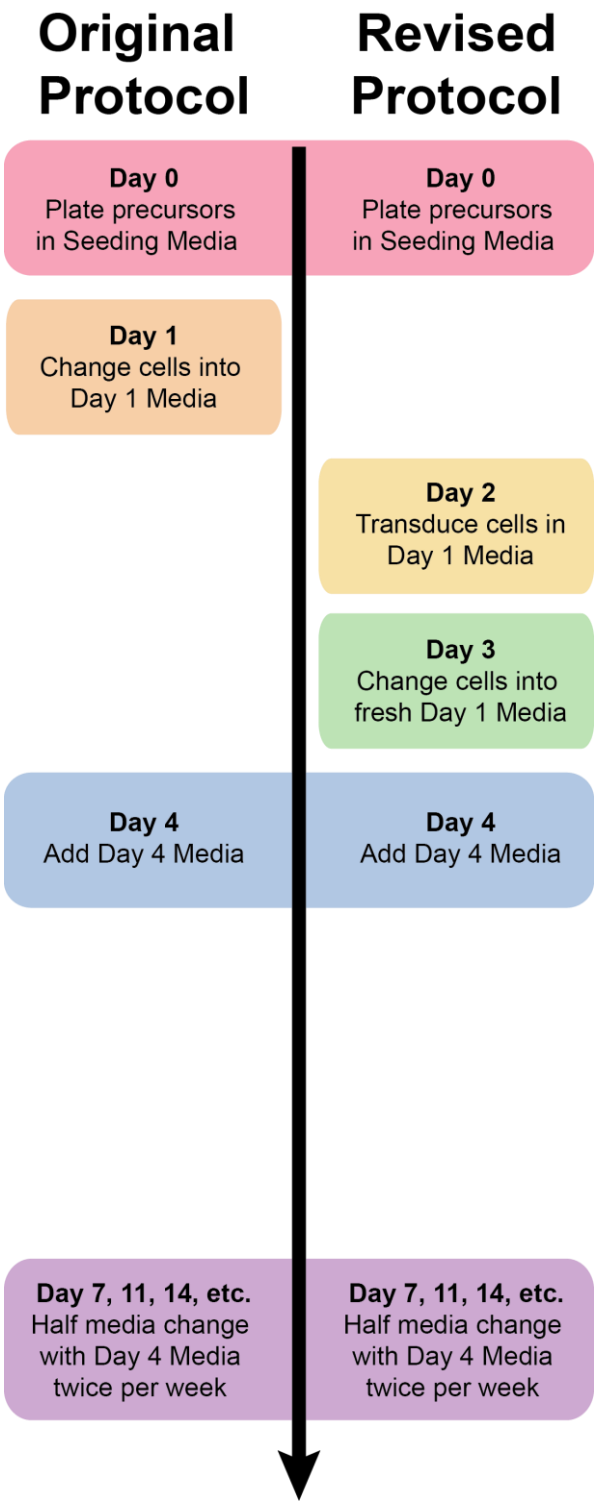


Figure C.1. Differentiation timeline. Comparison of the differentiation timeline described by BrainXell (left) versus the modifications made for the purposes of our screen (right).

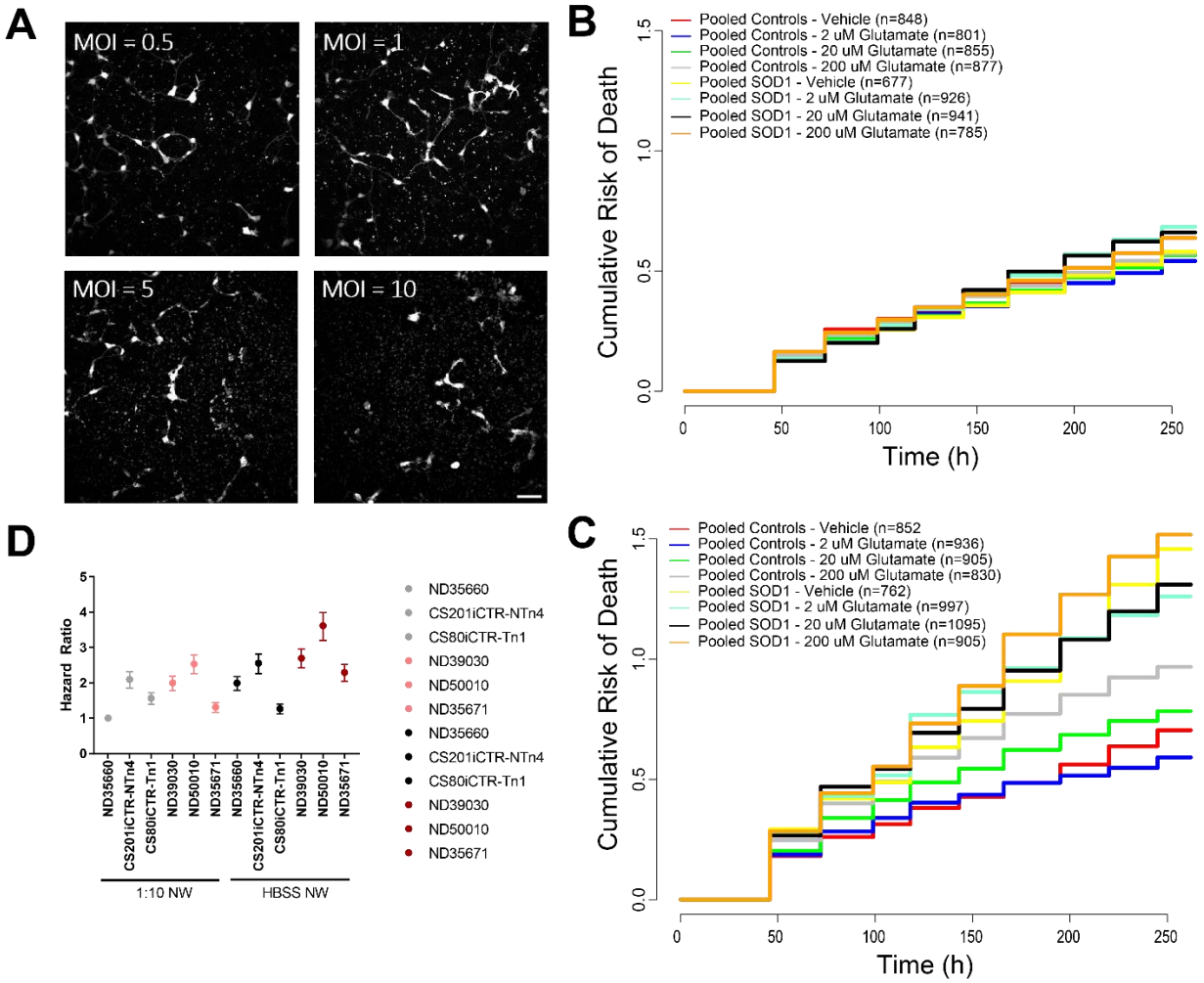


Figure C.2. Identification of a survival phenotype in SOD1-ALS iMNs. (A) To determine the MOI needed to fluorescently label the majority of cells, Day 2 iMNs were transduced with different amounts of virus. A MOI of 1 was sufficient to label >90% of plated cells. (B) Longitudinal fluorescence microscopy was used to track the relative survival of control or SOD1-ALS lines treated with various doses of glutamate. Over 10 days, there was remarkably little cell death even in the presence of an excitotoxic agent. (C) Glutamate in combination with complete replacement of culture media with HBSS resulted in gradual cell death over 10 days and revealed a disease-specific increase in the risk of death in SOD1-ALS lines. (D) The relative survival of individual control (gray) and SOD1-ALS (red) lines in either conditioned media diluted 1:10 in HBSS or complete nutrient withdrawal. Scale bar in (A), 50 μM.

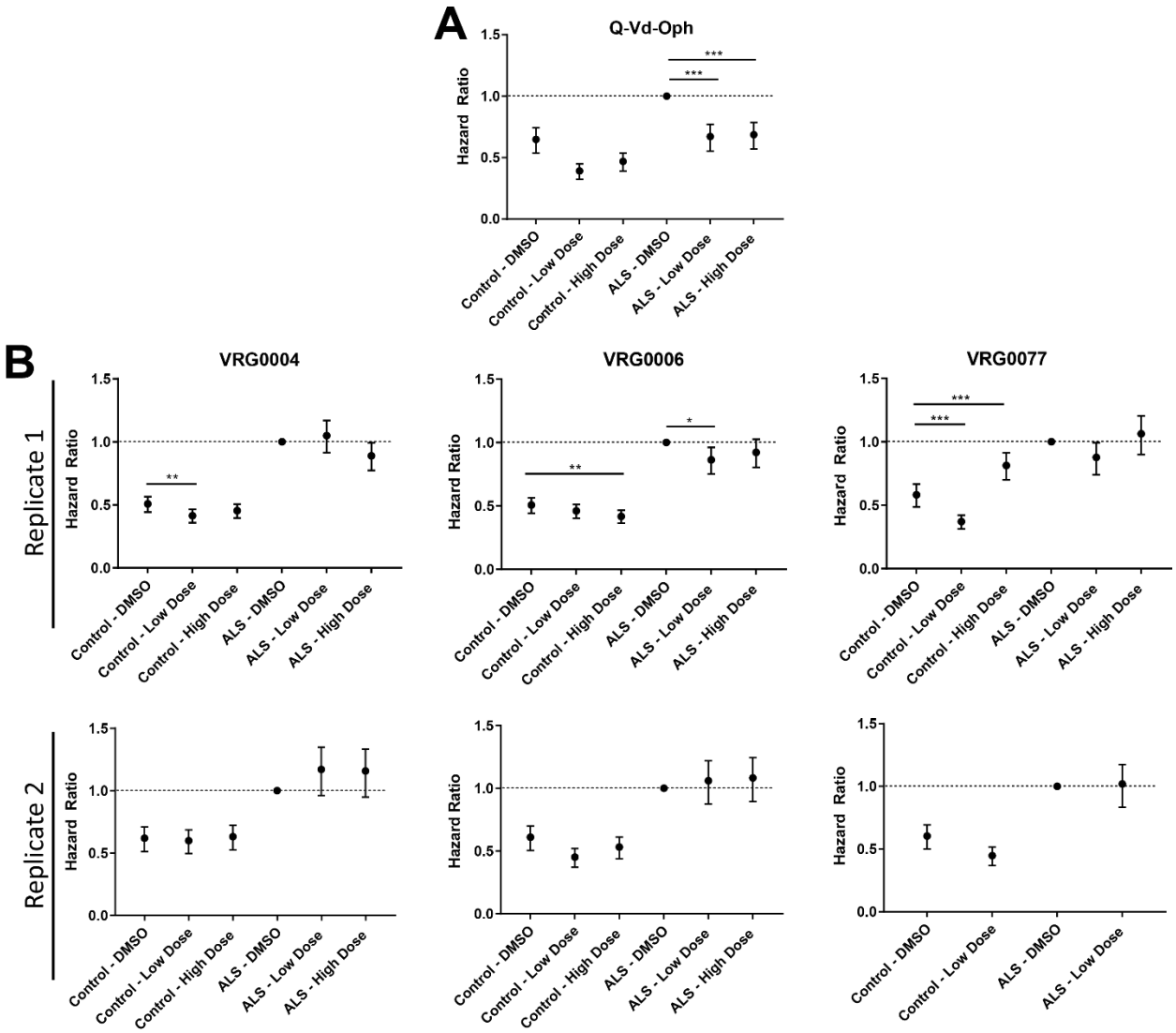


Figure C.3. Example plots for various compounds. (**A**) Treatment with the pan-caspase inhibitor Q-Vd-Oph significantly and consistently reduces cell death in SOD1-ALS iMNs compared to vehicle. (**B**) Example plots depicting the effects of compounds VRG004, VRG0006, and VRG0077. While some compounds show modest reduction in hazard ratio, these results are often difficult to reproduce.