JCI The Journal of Clinical Investigation

Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia

Mercedes Prudencio, … , Pietro Fratta, Leonard Petrucelli

J Clin Invest. 2020. https://doi.org/10.1172/JCI139741.

Research In-Press Preview **Neuroscience**

No treatment for frontotemporal dementia (FTD), the second most common early-onset dementia, is available but therapeutics are being investigated to target the two main proteins associated with FTD pathological subtypes: TDP-43 (FTLD-TDP) and tau (FTLD-tau). Testing potential therapies in clinical trials is hamstrung by our inability to distinguish between patients with FTLD-TDP and FTLD-tau. Therefore, we evaluated truncated stathmin-2 (*STMN2*) as a proxy of TDP-43 pathology, given reports that TDP-43 dysfunction causes truncated *STMN2* accumulation. Truncated *STMN2* accumulated in human iPSC-derived neurons depleted of TDP-43, but not in those with pathogenic *TARDBP* mutations in the absence of TDP-43 aggregation or loss of nuclear protein. In RNA-seq analyses of human brain samples from the NYGC ALS cohort, truncated *STMN2* RNA was confined to tissues and disease sub-types marked by TDP-43 inclusions. Lastly, we validated that truncated *STMN2* RNA is elevated in the frontal cortex of a cohort of FTLD-TDP cases but not in controls or cases with progressive supranuclear palsy (PSP), a type of FTLD-tau. Further, in FTLD-TDP, we observed significant associations of truncated *STMN2* RNA with phosphorylated TDP-43 levels and an earlier age of disease onset. Overall, our data uncovered truncated *STMN2* as a marker for TDP-43 dysfunction in FTD.

Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia

Mercedes Prudencio[†,1,2], Jack Humphrey[†,3,4], Sarah Pickles[†,1,2], Anna-Leigh Brown[5], Sarah E. Hill[6], Jennifer M. Kachergus[7], J. Shi[7], Michael G. Heckman[8], Matthew R. Spiegel[8], Casey Cook[1,2], Yuping Song[1], Mei Yue[1], Lillian M. Daughrity[1], Yari Carlomagno[1,2], Karen Jansen-West[1], Cristhoper Fernandez de Castro[1], Michael DeTure[1,2], Shunsuke Koga[1,2], Ying-Chih Wang[4], Prasanth Sivakumar[5], Cristian Bodo[5], Ana Candalija[9], Kevin Talbot[9], Bhuvaneish T. Selvaraj[10], Karen Burr[10], Siddharthan Chandran[10], Jia Newcombe[11], Tammaryn Lashley[12,13], Isabel Hubbard[14], Demetra Catalano[14], Duyang Kim[14], Nadia Propp[14], Samantha Fennessey[15], NYGC ALS Consortium[16], Delphine Fagegaltier[14], Hemali Phatnani[14,17], Maria Secrier[18], Elizabeth M. C. Fisher[5], Björn Oskarsson[19], Marka van Blitterswijk[1,2], Rosa Rademakers[1,2], Neil R. Graff-Radford[19], Bradley F. Boeve[20], David S. Knopman[20], Ronald C. Petersen[20], Keith A. Josephs[20], E. Aubrey Thompson[7], Towfique Raj[3,4,17], Michael Ward[6], Dennis W. Dickson[1,2], Tania F. Gendron[1,2], Pietro Fratta[5,17,*], Leonard Petrucelli[1,2,*].

†Equal contributions

Affiliations:

[1]Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA.

[2]Neuroscience Graduate Program, Mayo Clinic Graduate School of Biomedical Sciences, Jacksonville, FL, USA.

[3]Ronald M. Loeb Center for Alzheimer's Disease, Nash Family Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[4]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[5]Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, UK.

[6]National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA.

[7]Department of Cancer Biology, Mayo Clinic, Jacksonville, FL, USA.

[8]Division of Biomedical Statistics and Informatics, Mayo Clinic, Jacksonville, FL, USA.

[9]Nuffield Department of Clinical Neurosciences, University of Oxford, UK.

[10]UK Dementia Research Institute & Euan MacDonald Centre for MND Research, The University of Edinburgh, UK.

[11]NeuroResource, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, London, UK.

[12]Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology, London, UK.

[13]Queen Square Brain Bank for Neurological Disorders, Department of Clinical and Movement Neuroscience, UCL Queen Square Institute of Neurology, London, UK.

[14]Center for Genomics of Neurodegenerative Disease, New York Genome Center, New York, USA.

[15]New York Genome Center, New York, NY, USA.

[16]A complete list of members from the NYGC ALS Consortium can be found in the Supplemental Acknowledgments file.

[17]NYGC ALS Consortium.

[18]University College London Genetics Institute, London, UK.

[19]Department of Neurology, Mayo Clinic, Jacksonville, FL, USA.

[20]Department of Neurology, Mayo Clinic, Rochester, MN, USA.

*To whom correspondence should be addressed:

Leonard Petrucelli, Ph.D.

Department of Research, Neuroscience

Mayo Clinic College of Medicine

4500 San Pablo Rd

Jacksonville, FL 32224

Office: 1-904-953-2855

Fax: 1-904-953-6276

e-mail: petrucelli.leonard@mayo.edu

Pietro Fratta, MD, Ph.D.

Department of Neuromuscular Diseases

UCL Queen Square Institute of Neurology

Queen Square

London, WC1N 3BG

p.fratta@ucl.ac.uk

+44(0)2034484112

Abstract

No treatment for frontotemporal dementia (FTD), the second most common early-onset dementia, is available but therapeutics are being investigated to target the two main proteins associated with FTD pathological subtypes: TDP-43 (FTLD-TDP) and tau (FTLD-tau). Testing potential therapies in clinical trials is hamstrung by our inability to distinguish between patients with FTLD-TDP and FTLD-tau. Therefore, we evaluated truncated stathmin-2 (*STMN2*) as a proxy of TDP-43 pathology, given reports that TDP-43 dysfunction causes truncated *STMN2* accumulation. Truncated *STMN2* accumulated in human iPSC-derived neurons depleted of TDP-43, but not in those with pathogenic *TARDBP* mutations in the absence of TDP-43 aggregation or loss of nuclear protein. In RNA-seq analyses of human brain samples from the NYGC ALS cohort, truncated *STMN2* RNA was confined to tissues and disease sub-types marked by TDP-43 inclusions. Lastly, we validated that truncated *STMN2* RNA is elevated in the frontal cortex of a cohort of FTLD-TDP cases but not in controls or cases with progressive supranuclear palsy (PSP), a type of FTLD-tau. Further, in FTLD-TDP, we observed significant associations of truncated *STMN2* RNA with phosphorylated TDP-43 levels and an earlier age of disease onset. Overall, our data uncovered truncated *STMN2* as a marker for TDP-43 dysfunction in FTD.

Introduction

Tar DNA binding protein 43 (TDP-43) regulates multiple aspects of RNA metabolism, including RNA stabilization, transcription, splicing and transport (1). TDP-43 binds thousands of RNAs (2, 3) and, for particular genes, represses non-conserved cryptic exons in a species- (4) and cell-type specific manner (5, 6). Recently, two groups independently found that TDP-43 is required for the correct splicing and expression of stathmin-2 (*STMN2*) (7, 8), a microtubule-associated protein involved in axonal regeneration (9). As shown schematically in **Figure 1**, when TDP-43 is present and functional, it binds to GU-rich sequences in the first intron of the *STMN2* transcript, allowing normal splicing of intron 1 and repressing the usage of an alternative or cryptic polyadenylation site. Reduced TDP-43 binding to *STMN2* pre-mRNA, which could occur as a result of TDP-43 nuclear depletion, leads to the inclusion of a region within intron 1 (exon 2a) and introduces a premature stop codon and polyadenylation. This causes the accumulation of a truncated variant of *STMN2* that lacks exons 2 through 5.

Neuronal and glial inclusions of TDP-43 are a pathological hallmark of the majority (~97%) of amyotrophic lateral sclerosis (ALS) cases and of approximately half of cases with FTD (FTLD-TDP) (1). ALS is a motor neuron disease (MND) while FTD patients present clinically with one or more of the FTD syndromes characterized by behavior, executive function, language and motor impairments (10). TDP-43 abnormalities are proposed to play a direct role in these diseases since mutations in *TARDBP*, the gene that encodes TDP-43, were discovered in sporadic and familial ALS and FTD (11, 12). It is also notable that TDP-43 morphology and distribution associate with clinical phenotypes and neurodegeneration in key anatomical regions, such as frontal and temporal cortices in FTLD-TDP (13). Cells with TDP-43 pathology feature a redistribution of TDP-43 from the nucleus to the cytoplasm, where TDP-43 is ubiquitinated, hyperphosphorylated and cleaved into C-terminal fragments (14). Cases with FTLD-TDP, the neuropathological term to define FTD patients with TDP-43 pathology, can be further classified into different pathological subtypes based on the morphology and distribution of the TDP-43 inclusions observed (15, 16). These can take the form of neuronal cytoplasmic inclusions, dystrophic neurites, neuronal intranuclear inclusions, glial inclusions, perivascular inclusions, and as fine neurites in the hippocampus (14, 17-20). At present, five subtypes of FTLD-TDP (A-E) have been identified (15, 20-23), with types A-C accounting for more than 95% of all FTLD-TDP cases (24). FTLD-TDP type A, the most common, is characterized by the presence of all

inclusion types mentioned above, while types B-E are each characterized by the presence of a predominant inclusion type (19, 22). Furthermore, some FTLD-TDP subtypes are associated with particular mutations (20, 25, 26).

While ~50% of patients with an FTD syndrome have underlying TDP-43 pathology, such pathology can, at present, only be definitively confirmed at autopsy. Given efforts to develop therapies specific to FTLD pathological subtypes (i.e., FTD with TDP-43 vs. tau pathology), there is a clear need for biomarkers that discriminate between them. Further, biomarkers for diagnosis, prognosis or that can monitor TDP-43 activity in response to TDP-43-tailored therapies are also sorely lacking. While TDP-43 represents an attractive biomarker candidate to fulfill these needs, efforts to develop sufficiently sensitive immunoassays for the detection of TDP-43 or its pathological forms in biofluids, and that reliably differentiate between individuals with and without TDP-43 pathology have proven disappointing.(27) However, the nuclear clearance of TDP-43, in the absence of TDP-43 inclusions, has been associated with neuronal atrophy in FTLD-TDP and ALS-TDP cases (28), suggesting early mis-splicing events due to TDP-43 loss of function may represent a robust and early disease biomarker.

We thus propose that the mis-splicing of *STMN2* may provide a novel strategy to indirectly assess TDP-43 dysfunction. To test this hypothesis, we first validated the *STMN2* splicing event in iPSC-derived cortical and motor neurons that had TDP-43 depleted. Then we evaluated whether the truncated *STMN2* isoform accumulates in diseases and tissues characterized by TDP-43 pathology using datasets from RNA-seq studies of brain tissues from a large cohort of controls and of FTD and ALS cases with or without TDP-43 pathology. Furthermore, we measured truncated and full-length *STMN2* RNA levels in post-mortem medial frontal cortex tissue from FTLD-TDP, progressive nuclear palsy (PSP) and pathologically normal control brains. PSP was selected as disease control given that TDP-43 pathology is not found in the frontal cortex of FTLD-tau cases (29). We additionally evaluated whether truncated *STMN2* levels in the medial frontal cortex discriminated between FTLD-TDP cases and controls, and assessed associations of truncated *SMTN2* with phosphorylated TDP-43 (pTDP-43) and clinical features.

Results

TDP-43 depletion in human iPSC-derived neurons leads to aberrant STMN2 splicing

To confirm recent reports that correct *STMN2* splicing requires TDP-43 (7, 8), we used clustered regularly interspaced short palindromic repeat inactivation (CRISPRi) to knockdown TDP-43 expression in human induced pluripotent stem cells (iPSC)-derived cortical neurons that constitutively express nuclease deactivated Cas9 (dCas9) fused to blue fluorescent protein (BFP) and the transcriptional repressor Krüppel-associated box (KRAB) (dCas9-BFP-KRAB) (30). To do so, cells were transduced with lentivirus expressing either a single guide RNA (sgRNA) against the transcriptional start site of *TARDBP* (TDP-43 KD), or a control sgRNA predicted not to bind anywhere in the human genome (Controls). Silencing of *TARDBP* RNA was confirmed by qRT-PCR and RNA-seq (dataset labeled "a"), which showed knockdown efficiencies of TDP-43 of ~50% (47.4% for qRT-PCR, 63.9% for RNA-seq) (**Figure 2A**). Reduced *TARDBP* levels resulted in significant reductions of full-length *STMN2* RNA (56.7% for qRT-PCR, 65.5% for RNA-seq, **Figure 2B**), and a significant increase in truncated *STMN2* RNA [qRT-PCR (32.3-fold increase, **Figure 2C**), RNA-seq (**Figure 2D,** dataset labeled "a")] compared to levels in neurons expressing control sgRNA. In parallel, we analyzed published RNA-seq data from Klim *et al.* on iPSC-derived motor neurons in which *TARDBP* was knocked-down (7). Knockdown of *TARDBP* (**Supplementary Figure 1A**) was confirmed in the Klim dataset (dataset labeled "b") and, as expected, accompanied by reduced full-length *STMN2* RNA (**Supplementary Figure 1B**) and increased truncated *STMN2* RNA (**Figure 2D**, dataset labeled "b").

We also investigated whether *TARDBP* disease-causing mutations, in the absence of TDP-43 protein mislocalization or reduction of nuclear levels, are sufficient to impact *STMN2* splicing. We evaluated truncated *STMN2* expression using RNA-seq data from two independent studies using iPSC-derived motor neurons from patients carrying ALS pathogenic *TARDBP* mutations and controls (**Supplementary Table 1**). In contrast to neurons in which *TARDBP* was downregulated, no increase in truncated *STMN2* was observed in motor neurons with *TARDBP* mutations (**Figure 2D**, datasets labeled "c" and "d"). This observation is consistent with the lack of change in TDP-43 localization (**Figure 2E**), *TARDBP* RNA levels (**Supplementary Figure 2**) and TDP-43 protein levels (**Supplementary Figure 3**) in *TARDBP* mutant motor neurons. Overall, these data suggest that truncated *STMN2* accumulation in iPSC-derived neurons results from a loss of TDP-43 function as a result of TDP-43 depletion and not from splicing function

changes arising from mutations in *TARDBP* (31). The redistribution of TDP-43 from the nucleus and its sequestration into cytoplasmic inclusions, a pathological hallmark of FTD and ALS, with and without a genetic link to *TARDBP,* are consistent with a loss of function mechanisms that would lead to *STMN2* mis-splicing and an upregulation of truncated *STMN2* expression and prompted us to further examine this in human tissue from FTD/ALS patients.

Analyses of bulk brain tissue RNA-seq data uncover tissue-specific truncated STMN2 RNA accumulation in diseases characterized by TDP-43 pathology

We quantified the expression of full-length and truncated *STMN2* in the New York Genome Center (NYGC) ALS Consortium RNA-seq dataset across multiple neuroanatomical regions and diseases. As of March 2020, 11 tissues could be examined, with 1,659 RNA-seq samples from 439 donors matching established criteria for inclusion (see **Methods**). Donors were diagnosed as non-neurological disease controls (controls), FTD, ALS, FTD with ALS (ALS/FTLD) or ALS with suspected Alzheimer's disease (ALS/AD). FTD cases were classified based on their neuropathological diagnosis: FTLD-TDP, FTLD-tau, and FTLD-FUS. ALS samples were divided into those with a SOD1 mutation (ALS-SOD1), from those without a SOD1 mutation and thus assumed to have TDP-43 pathology (ALS-TDP). However, pathology confirming the presence of TDP-43 mislocalization in ALS-TDP cases was not systematically reported (see **Methods**).

For each sample, full-length *STMN2* expression was estimated using all RNA-seq reads covering known *STMN2* exons, whereas the truncated *STMN2* isoform was quantified using high confidence spliced RNA-seq reads that unambiguously connect *STMN2* exon 1 with the cryptic exon 2a (exon 1-exon 2a, **Figure 3A**). Full-length *STMN2* was highly expressed in both cortical and cerebellar tissues (median transcripts per million (TPM) in frontal cortex = 58.0, and cerebellum = 73.1) compared to spinal cord (median TPM in lumbar spinal cord = 6.7), presumably due to its restricted expression to neurons and the higher white matter composition in the spinal cord (**Supplementary Figure 4**).

We then estimated the expression of truncated *STMN2* in the NYGC ALS cohort in disease groups across tissues. At least two reads spanning exon 1-exon 2a junction could be detected in 398 of the 1,659 samples in the dataset (24.0%) (**Supplementary Figure 5**). Samples that expressed the splice junction did so at a median read count of six (0.13 reads per million)

and a maximum read count of 120 (2.5 reads per million) (**Supplementary Figure 6**). We therefore chose to model truncated *STMN2* expression as a binary outcome, with a detection threshold of two junction reads uniquely mapped across the intron boundary.

Truncated *STMN2* was present in disease subtypes known to have TDP-43 pathology, being detected in the frontal cortex of 82.4% of FTLD-TDP cases but absent in FTLD-FUS, FTLD-tau, and controls (**Figure 3B)**. In ALS lumbar spinal cord, truncated *STMN2* was detected in 62.2% of ALS cases with presumed TDP-43 pathology (ALS-TDP) but not in ALS-SOD1 cases or control individuals (**Figure 3C**).

We observed an exquisite specificity of truncated *STMN2* expression in tissues known to be affected with TDP-43 pathology. In FTLD-TDP, where frontal cortex, temporal cortex and cerebellum datasets were available, truncated *STMN2* was detected in frontal and temporal cortices but not in cerebellum (**Figure 3D**). In ALS-TDP, which also included hippocampus, temporal cortex, motor cortex, occipital cortex and spinal cord samples, truncated *STMN2* expression was most prevalent in the spinal cord, followed by the motor cortex and frontal cortex, and with few to zero cases in the occipital cortex and cerebellum, respectively (**Figure 3E, Supplementary Figure 5**). As anticipated, patients with concurrent FTD and ALS (ALS/FTLD) had truncated *STMN2* in frontal and temporal cortices, as well as motor cortex and spinal cord (**Supplementary Figure 7A**). In contrast, patients with ALS/AD strongly resembled ALS-TDP patients in their truncated *STMN2* detection patterns with most truncated *STMN2* found in spinal cord and reduced or no expression in cortical brain regions (**Supplementary Figure 7B**).

Together these findings demonstrate that truncated *STMN2* RNA is detected in diseases and tissues characterized with TDP-43 pathology, supporting its potential as a marker for TDP-43 dysfunction.

Truncated STMN2 RNA is elevated in FTLD-TDP frontal cortex

Given that truncated *STMN2* RNA was detected in FTLD-TDP cases, but not in FTLD-FUS or FTLD-tau cases (**Figure 3B**), this aberrant splicing event may provide a means to identify FTD patients with marked TDP-43 pathology. To further examine the relationship between TDP-43 and *STMN2* in FTD, we used the NanoString Plexset platform to systematically quantify truncated and full-length *STMN2* RNA levels in frontal cortex tissue from a cohort of

238 FTD patients with immunohistologically-confirmed TDP-43 immunoreactive inclusions, a cohort of 33 cognitively normal controls that were TDP-43 negative (henceforth referred to as controls), and 41 PSP cases (**Table 1**). The latter are characterized by tau pathology and neurodegeneration in the frontal cortex.

As mentioned, FTLD-TDP can be further classified based on the morphology and distribution of TDP-43 inclusions. As such, our cohort of FTLD-TDP cases can be classified into different TDP-43 pathological subtypes: 117 were type A (27 and 34 carried a mutation in *C9orf72* or *GRN*, respectively), 66 type B (21 and 1 carried a mutation in *C9orf72* or *GRN*, respectively), 43 type C (2 and 1 carried a mutation in *C9orf72* or *GRN*, respectively), and two type D (both *VCP* mutants) (**Table 1**). In addition to conducting our analyses on all FTLD-TDP cases combined, we also evaluated whether truncated *STMN2* accumulation preferentially associated with a particular TDP-43 subtype. In the frontal cortex, truncated *STMN2* RNA was significantly elevated in all FTLD-TDP cases compared to controls or PSP disease controls in both unadjusted analysis ($P < 0.001$; **Figure 4A**; **Table 2** and **Supplementary Table 2**), and analysis adjusted for age at death, sex, RIN ($P = 0.009$ vs. controls, $P < 0.001$ *vs*. PSP; **Figure 4A**; **Table 2** and **Supplementary Table 2**). We also noted that the levels of truncated *STMN2* in PSP cases were significantly lower compared to controls ($P = 0.028$, adjusted analyses for age, sex and RIN), potentially due to the neuronal loss in PSP cases.

As in all FTLD-TDP cases combined, truncated *STMN2* RNA levels in the frontal cortex were increased in FTLD-TDP for each TDP-43 type (A-D) compared to PSP cases in both unadjusted and adjusted analyses (**Table 2** and **Supplementary Table 2)**. Compared to controls, truncated *STMN2* levels were significantly increased in FTLD-TDP types A, B and C, but not D, in unadjusted analyses (**Table 2** and **Supplementary Table 2**) although only type A remained significant in analyses adjusted for age, sex and RIN (**Table 2** and **Supplementary Table 2**).

We observed that truncated *STMN2* levels in the frontal cortex discriminated between the FTLD-TDP cases and PSP controls with an area under the receiver operating characteristic curve (AUC) of 0.76 (95% CI = 0.69-0.84) (**Table 2**). Frontal cortex truncated *STMN2* also discriminated between FTLD-TDP cases and cognitively normal controls with a more modest AUC of 0.67 (95% CI = 0.57-0.78) (**Supplementary Table 2**). Furthermore, truncated *STMN2* levels similarly discriminated between controls (cognitively normal or PSP disease controls) and FTLD-TDP patients from all TDP-43 types, which was most prominent for types A, C and D

(**Table 2** and **Supplementary Table 2**). In contrast, truncated *STMN2* levels did not effectively in distinguish PSP patients from controls (AUC: 0.60, 95% CI = 0.47-0.73).

The increase in truncated *STMN2* in FTLD-TDP frontal cortex was paired with a decrease in full-length *STMN2* ($P < 0.001$, **Figure 4B, Supplementary Table 3**). Of interest, full-length *STMN2* was also decreased in PSP cases compared to controls ($P < 0.001$, Figure **4B, Supplementary Table 3**), likely due to neurodegeneration in this region. Full-length *STMN2* levels also discriminated controls from FTLD-TDP (AUC: 0.74) and PSP (AUC: 0.83), but were not as effective in distinguishing FTLD-TDP from PSP cases (AUC: 0.60, 95% CI = 0.51-0.68; **Supplementary Table 3**).

Taken together, our data show that truncated *STMN2* RNA levels are significantly elevated in FTLD-TDP cases compared to controls, and may be used to discriminate FTD patients with TDP-43 pathology from both cognitively normal and disease controls.

Truncated STMN2 RNA associates with pTDP-43 burden in the frontal cortex of FTLD-TDP cases

That truncated *STMN2* was elevated in the frontal cortex of FTLD-TDP but not in PSP cases (**Figure 4A**) or other forms of FTLD (FTLD-tau, FTLD-FUS) or ALS (ALS-SOD1) (**Figure 3**), suggests that truncated *STMN2* production is caused by the loss of TDP-43 nuclear function resulting from its mislocalization to the cytoplasm and/or its sequestration into inclusions – hallmark features of FTLD-TDP. To evaluate this possibility further, we examined whether full-length or truncated *STMN2* RNA in the frontal cortex correlated with pTDP-43 accumulation. To this end, we prepared urea-soluble protein lysates from the same tissues used to measure *STMN2* isoforms, and quantified pTDP-43 in these lysates using a previously described immunoassay.(32-34) While the levels of full-length *STMN2* did not associate with pTDP-43 levels (β: 0.02, $P = 0.53$, **Supplementary Table 4**), higher pTDP-43 levels were associated with higher truncated *STMN2* RNA (**Figure 4C**) in both unadjusted analyses (β: 0.20, $P < 0.001$, **Table 3**) and in a multivariable analysis adjusting for age, sex and TDP-43 subtype (β: 0.19, $P < 0.001$, **Table 3**). Taken together, our data suggest that truncated *STMN2* RNA, but not full-length *STMN2*, accumulates in brain regions marked by TDP-43 pathology, and associates with the burden of pTDP-43.

Truncated STMN2 RNA levels associate with earlier age of FTLD-TDP onset

We additionally examined whether truncated or full-length *STMN2* levels in the frontal cortex associate with clinical characteristics of individuals with FTLD-TDP. We found a significant inverse correlation between truncated *STMN2* RNA and age at disease onset (**Figure 4D**) in analyses adjusting for sex, age at death and TDP-43 type and after correcting for multiple testing (P = 0.005, **Table 3**). However, survival after onset did not associate with truncated *STMN2* RNA, correcting for age at onset, sex and TDP-43 subtype (P = 0.32, **Table 3**). Of interest, we observed a nominal but not statistically significant association between lower truncated *STMN2* levels in males (β: -0.35, P = 0.034 in adjusted analyses for age and TDP-43 type, **Supplementary Table 4**). In contrast, full-length *STMN2* levels failed to associate with clinical features of FTLD-TDP, including age of onset and survival (**Supplementary Table 4**).

Discussion

FTLD, the neuropathological substrate of FTD, is the second most common cause of early onset dementia (35). While no effective treatment is available, several therapeutic approaches are being investigated to target TDP-43 and tau, the two proteins most commonly associated with the major pathological FTLD subtypes. Their testing in clinical trials, however, faces a considerable barrier: our inability to determine the underlying protein pathology in living patients with FTD. While the pathomechanisms driven by TDP-43 and tau differ, the clinical symptoms of FTD patients with TDP-43 or tau pathology overlap. Ultimately, this hinders the selection of sporadic FTD patients suitable for a particular clinical trial. Since there is a clear need for biomarkers to identify FTD patients with TDP-43 pathology, we evaluated whether truncated *STMN2* could serve as a proxy for TDP-43 pathology. We show, by RNA-seq and RNA quantification techniques that truncated *STMN2* RNA levels were elevated in human iPSC-derived neurons depleted of TDP-43 and in bulk tissue from FTD and ALS patients with TDP-43 pathology. Truncated *STMN2* was not detected in iPSC-derived neurons with pathogenic *TARDBP* mutations where no change in TDP-43 levels, localization or TDP-43 aggregates are detected, supporting the fact that loss of TDP-43 nuclear function, and not splicing changes induced by pathogenic mutations are driving this event. We were unable to assess post mortem tissue from patients carrying *TARDBP* mutations, which are rare in the FTD and ALS populations, as a group, but given the fact these patients develop florid TDP-43 pathology, we

expect truncated *STMN2* to be present similarly to sporadic ALS and FTLD-TDP cases. We additionally found that truncated *STMN2* RNA accumulated in the frontal cortex of FTLD-TDP cases but not in controls or PSP cases, and significantly correlated with pTDP-43 levels and an earlier age of disease onset. Taken together, our data support the notion that truncated *STMN2* production results from TDP-43 loss of function.

As mentioned above, two independent reports recently revealed that TDP-43 regulates *STMN2* expression, and that TDP-43 depletion produces a truncated, non-functional *STMN2* variant (7, 8). Increased truncated *STMN2* RNA was detected in spinal cord tissue of 17 ALS patients but not of six controls, and *in situ* hybridization studies found truncated *STMN2* RNA in lower and upper motor neurons in five ALS patients (8). We thus evaluated truncated *STMN2* as a proxy TDP-43 activity marker for FTD using post-mortem RNA-seq data across 11 tissues available from a large cohort of FTD, ALS and FTLD-ALS cases. Notably, our RNA-seq analyses detected the *STMN2* cryptic exon that gives rise to truncated *STMN2* only in FTD and ALS cases with TDP-43 pathology. In addition, the presence of abnormal *STMN2* was highly specific to brain regions characterized by TDP-43 pathology.

In addition to our analyses of RNA-seq data, we directly measured truncated *STMN2* RNA in post-mortem frontal cortex tissue from FTLD-TDP cases, which again revealed an aberrant accumulation of truncated *STMN2* in comparison to controls and PSP cases. In fact, truncated *STMN2* RNA levels were found to discriminate between FTLD-TDP and PSP cases with an area under the ROC curve of 0.76. As anticipated, truncated *STMN2* accumulation in FTLD-TDP cases was accompanied by a decrease in full-length *STMN2*, as was observed in cell culture models (7, 8). Full-length *STMN2* was also decreased in PSP cases, which were confirmed to lack TDP-43 pathology. Given the fact that *STMN2* is expressed by neurons and its expression is abundant in the frontal cortex, the decreased levels of full-length *STMN2* in PSP cases likely reflect neurodegenerative neuronal loss. Indeed, a recent report found full-length *STMN2* RNA levels were also decreased in the brain of Parkinson's disease patients (36), possibly arguing for the use of *STMN2* levels as a neurodegenerative disease biomarker. It is also possible that, aside from TDP-43, various other factors regulate full-length *STMN2* expression. This could explain why full-length *STMN2* levels did not correlate with pTDP-43 burden or with clinical features.

The accumulation of truncated *STMN2* RNA in the frontal cortex of FTLD-TDP but not PSP cases supports truncated *STMN2* production being caused by pathological changes in TDP-43. In line with this, we observed a significant association of truncated *STMN2* RNA with pTDP-43 protein, which remained statistically significant after controlling for multiple variables (age, sex, and TDP-43 subtype). However, it is possible that pTDP-43 levels alone may not be a perfect corollary for nuclear TDP-43 depletion. Of interest, the levels of truncated *STMN2* RNA did not differ among different TDP-43 subtypes, suggesting it similarly accumulates in response to different types of TDP-43 pathology.

Importantly, higher truncated *STMN2* RNA levels in the frontal cortex at death significantly associated with an earlier age at disease onset, but we found no association with survival after disease onset. Given that truncated *STMN2* RNA serves as a proxy for TDP-43 dysfunction, this earlier age of onset in patients with higher post-mortem truncated *STMN2* RNA levels could point to more severe defects in RNA metabolism resulting from aberrant TDP-43 function, including loss of *STMN2* function.

Strengths of our study include the use of a clinically well-characterized, immunohistologically-confirmed cohort of 238 FTLD-TDP cases from which high integrity RNA was extracted from the frontal cortex, and the use of quantitative measures of truncated *STMN2* RNA and pTDP-43 protein. Moreover, we evaluated multiple neuroanatomical regions in bulk tissue RNA-seq from a large dataset that included 1,659 FTD and ALS cases with and without TDP-43 pathology. Further, our study also uses three different approaches (RNA-seq, NanoString and qRT-PCR, the latter also used to validate truncated *STMN2* in FTLD-TDP cases as shown in **Supplementary Figure 9**), which have different sensitivity and specificity linked to the methodology, but all detect significant differences in truncated *STMN2*, highlighting the reproducibility of the finding. Our study also has some limitations. While we specifically designed our study to evaluate truncated *STMN2* as a marker of TDP-43 pathology – information important to establish its value as a candidate biomarker – we did not measure the translation product from this truncated variant due to a lack of commercially available antibodies. At present, it is unclear whether the predicted peptide of only 16 amino acids is translated from truncated *STMN2* RNA. Nevertheless, our findings indicate that the development of novel tools (antibodies, immunoassays) for the detection of truncated *STMN2* protein, and additional studies to evaluate the role of truncated *STMN2* RNA and protein in disease pathogenesis and as

biomarkers, are well merited. Critical next steps to hasten the translation of this finding to the clinic include determining whether truncated *STMN2* RNA or protein are detected in FTD patient biofluids (e.g., spinal fluid and blood) and, if so, whether they discriminate between FTD subtypes (i.e., with or without TDP-43). While expression of truncated *STMN2* in affected post-mortem tissues can discriminate FTLD-TDP from PSP cases, which has important implications as it supports truncated STMN2 as a viable proxy for the presence of TDP-43 pathology, the ability to discriminate FTLD-TDP from PSP cases was only modest (AUC: 0.76). RNA degradation invariably occurs in post-mortem brain tissues, possibly contributing to this modest AUC, and potentially pointing toward the possibility of detecting truncated *STMN2* in CSF or blood, which are typically frozen immediately after collection, to discriminate between patient groups. Nonetheless, sensitive methods will be needed to detect truncated *STMN2* RNA and/or protein in human biofluids. Additionally, determining at what point during the course of disease truncated *STMN2* levels begin to be upregulated will have to be evaluated, likely requiring longitudinal collections of CSF and blood from, first from ALS cases, who are almost certain to develop TDP-43 pathology and then from FTD and FTD/ALS individuals harboring genetic variants associated with TDP-43 dysfunction.

If, as anticipated, truncated *STMN2* RNA or protein are detected in biofluids of FTD patients with TDP-43 pathology, it may have the potential to similarly identify those patients with Alzheimer's disease (AD) that have TDP-43 pathology. Indeed, up to 67% of patients with AD develop TDP-43 inclusions and, compared to AD patients without TDP-43 pathology, they are cognitively and functionally worse, suggesting that abnormal TDP-43 causes a modified AD phenotype (37-46). Our work also highlights how future means to detect truncated *STMN2* RNA or protein in biofluids may provide important information on disease activity, and therefore be useful in assessing drug efficacy in clinical trials. Thus, these endeavors are expected to increase the likelihood of success in developing effective therapeutics and improving the care of patients with TDP-43 proteinopathies.

Methods

RNA extraction, complementary DNA synthesis and qRT-PCR in iPSC-derived neurons

iPSC-derived neurons were harvested and cell pellets were lysed in TRI reagent and then total RNA was extracted as per the manufacturer's instructions using the Direct-zol RNA

MiniPrep kit (Zymo Research), and 850 ng of RNA was used for reverse transcription polymerase chain reaction (RT-PCR). RT-PCR of RNA samples used the High Capacity cDNA Transcription Kit with random primers (Applied Biosystems), and quantitative real-time PCR (qRT-PCR) was performed using SYBR GreenER qPCR SuperMix (Invitrogen). Samples were run in triplicate, and qRT-PCRs were run on a Prism 7900HT Fast Real-Time system (Applied Biosystems). Relative quantification was determined using the ΔΔCt method and normalized to the endogenous controls *RPLP0* and *GAPDH*. List of primers and their sequences: *STMN2* forward 5'-AGCTGTCCATGCTGTCACTG-3', *STMN2* reverse 5'-GGTGGCTTCAAGATCAGCTC-3', Truncated *STMN2* forward 5'-GGACTCGGCAGAAGACCTTC-3', Truncated *STMN2* reverse 5'-GCAGGCTGTCTGTCTCTCTC-3', *TARDBP* forward 5'-AATTCTGCATGCCCCAGA-3', *TARDBP* reverse 5'-GAAGCATCTGTCTCATCCATTTT-3', *GAPDH* forward 5'-GTTCGACAGTCAGCCGCATC-3', *GAPDH* reverse 5'-GGAATTTGCCATGGGTGGA-3', *RPLP0* forward 5'-TCTACAACCCTGAAGTGCTTGAT-3', *RPLP0* reverse 5'-CAATCTGCAGACAGACACTGG-3'.

RNA-seq analyses from iPSC-derived neurons

Data from Klim *et al* (7), was downloaded from Gene Expression Omnibus (GEO) accession GSE121569. All iPSC cells lines were analyzed with the same analysis pipeline. Raw FASTQ files were trimmed using fastp (47) and aligned to the hg38 build (GRCh38.primary_assembly) of the human reference genome using STAR (2.7.2a) (48). Gene abundance for *TARDBP* and *STMN2* was calculated using featureCounts (49) using gene models from GENCODE GTF v31 and converted to reads per kilobase per million (RPKM) using edgeR (50). The relative abundance of truncated *STMN2* as a proportion of total *STMN2* was expressed in terms of percent spliced in (PSI) (51). This was calculated by parsing the splice junction tables from STAR and taken as the percentage of all uniquely mapped *STMN2* junction reads from exon 1 to exon 2a over the sum of all junction reads coming from exon 1.

NYGC ALS Consortium cohort

FTD cases were divided according to a pathologist's diagnosis of FTD with TDP-43 inclusions (FTLD-TDP), Tau (FTLD-tau) or FUS inclusions (FTLD-FUS). ALS samples were

divided into the following subcategories using the available Consortium metadata: ALS with or without reported SOD1 mutations (ALS-TDP and ALS-SOD1); ALS with frontotemporal dementia (ALS/FTLD); and ALS with Alzheimer's disease (ALS/AD). All non-SOD1 ALS samples were grouped as 'ALS-TDP' in this work for simplicity, although reporting of post-mortem TDP-43 inclusions was not systematic and therefore not integrated in the metadata. Confirmed TDP-43 pathology post-mortem was reported for all FTLD-TDP samples.

NYGC ALS Consortium RNA-seq library preparation from bulk tissue RNA

Consortium sample processing has been, in part, previously described (52). In brief, RNA was extracted from flash-frozen post-mortem tissue using Trizol (ThermoFisher)-chloroform, followed by column purification (QIAGEN RNeasy minikit). RIN was assessed on a Bioanalyzer (Agilent). RNA-seq libraries were prepared from 500 ng of total RNA using KAPA Stranded RNA-Seq Kit with RiboErase (KAPA Biosystems) for rRNA depletion and Illumina-compatible indexes (NEXTflex RNA-seq Barcodes, NOVA-512915, PerkinElmer; and IDT for Illumina TruSeq UD Indexes, 20022370). Pooled libraries (average insert size: 375 bp) passing quality criteria were sequenced either on an Illumina HiSeq 2500 (125 bp paired end) or an Illumina NovaSeq (100 bp paired end). Samples are subject to extensive sequencing and RNA-seq quality control metrics performed at NYGC that will be described elsewhere. Notably, a set of >250 markers is used to confirm tissue, neuroanatomical regions and sex in RNA-seq data. Samples passing these metrics are available for distribution. Samples had a median sequencing depth of 42 million read pairs, with a range of between 16 and 167 million.

Analyses of bulk tissue RNA-seq data

Samples were uniformly processed using RAPiD-nf (53), an efficient RNA-seq processing pipeline implemented in the NextFlow framework (54). Following adapter trimming with trimmomatic (v0.36) (55), all samples were aligned to the hg38 build (GRCh38.primary_assembly) of the human reference genome using STAR (2.7.2a) (48) with indexes created from GENCODE (v30) (56). Gene expression was quantified using RSEM (1.3.1) (57). Quality control (QC) was performed using Samtools (58), PicardTools (59), and collated together using MultiQC (60). Exploratory plots were created in R (v3.6.0) (61) using dplyr, tidyr, stringr, ggplot2, patchwork and ggpubr packages (62-64).

RNA-seq samples were subject to quality control modelled on the Genotype Tissue Expression Consortium (65). Any sample failing one of the following sequencing metric thresholds was removed: unique alignment rate < 90%; ribosomal bases > 10%; mismatch rate > 1%; duplication rate > 0.5%; intergenic bases < 10.5%; ribosomal bases > 0.1%. Sex was verified using XIST and UTY expression. Tissue identity was confirmed using the expression of cerebellar marker *CBNL1* and cortical marker *NRGN*. Spinal cord samples were confirmed by expression of the oligodendrocyte marker *MBOP*. Following QC, 1,905 samples of 1,924 were retained for analysis. When comparing subgroups and tissues, only tissues with at least 50 samples and disease subgroups with at least four individuals were kept for analysis. This resulted in 1,659 samples from 11 tissues and 439 donors.

Expression of the full-length *STMN2* isoform was estimated using RSEM, summing the estimates of the three annotated isoforms in GENCODE v30 into a gene-level estimate of expression normalized in terms of transcripts per million (TPM). Due to the large size of the NYGC ALS cohort, RNA-seq data was produced on two different sequencing platforms. Exploratory analyses showed a substantial batch effect in gene expression due to the two sequencing platforms. We therefore split the dataset by sequencing platform, along with tissue and disease status. Comparing ALS-TDP and non-neurological controls, *STMN2* had an inconsistent pattern of differential expression with no tissue showing a marked difference in both platforms (**Supplementary Figure 4A**). In addition, RNA degradation is known to affect estimation of transcript abundance (66, 67) and we reasoned that systematic differences in RNA quality amongst samples sequenced on different platforms, and between disease and control groups or tissues, could confound our estimation of *STMN2* expression. We compared the RIN values for each sample between ALS-TDP and control groups and found significant differences in RIN values in some comparisons (**Supplementary Figure 4B)**. The small number of controls available limited further comparisons in some tissues. We then correlated RIN with *STMN2* expression and observed a positive correlation in multiple tissues and both sequencing platforms (**Supplementary Figure 4C**).

Quantifying STMN2 splicing from RNA-seq data

For downstream analysis of *STMN2* splice junction reads, a bioinformatic pipeline was created using Snakemake (68, 69). Uniquely mapping reads within the *STMN2* locus were

extracted from each sample using Samtools. Any read marked as a PCR duplicate by PicardTools was discarded. Splice junction reads were then extracted with Regtools (70), using a minimum 8 bp as an anchor on each side of the junction and a maximum intron size of 500 kb. Junctions from each sample were then clustered together using Leafcutter (71) with relaxed junction filtering (minimum total reads per junction = 30, minimum fraction of total cluster reads = 0.0001). This produced a matrix of junction counts across all samples. Using coordinates from the Klim (7) and Melamed (8) publications, the TDP-43-associated junction of the novel truncated *STMN2* isoform was extracted from the matrix. A detection threshold of two or more uniquely mapping reads was applied. We found no association between truncated *STMN2* detection and RIN, sequencing platform or total library size (**Supplementary Figure 8A-C**). The relationship between full length and truncated *STMN2* expression is complex with no overall direction of association across tissues (**Supplementary Figure 8D**).

RNA extraction and NanoString analyses in post-mortem tissues

Using the RNAeasy Plus Mini Kit (Qiagen), RNA was extracted from frozen human post-mortem frontal cortex tissues (33). RNA with RIN $\geq$ 7.0, as determined using an Agilent 2100 bioanalyzer (Agilent Technologies), was used to measure transcript levels using NanoString Plexset platform (250 ng human brain RNA, as per manufacturer's instructions). Transcript level abundance was determined using nSolver 4.0 analyses software (NanoString technologies) and normalized to *HTRP1,* a housekeeping gene that was unaltered across disease subtypes in our dataset, and previously used to normalize human brain transcripts (72). Sequences of the NanoString probes used in this study are included in **Supplementary Table 5**.

Phosphorylated TDP-43 immunoassay

The phosphorylated TDP-43 immunoassay was performed as previously described (33). Briefly, approximately 50 mg of postmortem tissue from the frontal cortex of patients or unaffected individuals were homogenized in cold RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% sodium deoxycholate, 1% Nonidet P-40, 0.1% sodium dodecyl sulfate, protease and phosphatase inhibitors) and sonicated on ice. Homogenates were centrifuged at 100,000 x g for 30 minutes at 4°C. The supernatant was collected, and the pellet was resuspended in RIPA buffer, sonicated and centrifuged again to prevent carry-over of soluble material. The RIPA-

insoluble pellet was extracted using 7 M urea buffer, sonicated and centrifuged at 100,000 x g for 30 minutes at 22°C. Protein concentration of the urea-soluble supernatant was determined by Bradford assay. Phosphorylated TDP-43 levels were evaluated using a sandwich Meso Scale Discovery (MSD) immunoassay. The capture antibody was a mouse monoclonal antibody that detects TDP-43 phosphorylated at serines 409/410 (1:500, #CAC-TIP-PTD-M01, Cosmo Bio U.S.A) and the detection antibody was a Sulfo-tagged rabbit polyclonal C-terminal TDP-43 antibody (2 μg/ml, 12892-1-AP, Proteintech). Urea-soluble fractions were diluted in TBS to 35 μg of protein per well and tested in duplicate wells. Response values corresponding to the intensity of emitted light upon electrochemical stimulation of the assay plate using the MSD QUICKPLEX SQ120 were acquired.

Statistics

Comparisons of truncated *STMN2* RNA in the frontal cortex between the separate non-FTLD-TDP groups of controls and PSP cases and various groups of FTLD-TDP patients (all FTLD-TDP patients as well as four different subgroups according to TDP-43 subtype) were made using single-variable (i.e. unadjusted) and multivariable linear regression models. Multivariable models were adjusted for age at death, sex, and RIN, and truncated *STMN2* RNA in the frontal cortex was examined on the base-2 logarithm scale owing to its skewed distribution. Regression coefficients (referred to as β) and 95% confidence intervals (CIs) were estimated, and are interpreted as the difference in mean truncated *STMN2* RNA level in the frontal cortex (on the base 2 logarithm scale) between the given group of FTLD-TDP patients and the given reference group (cognitively normal controls or PSP cases). Additionally, to further assess the ability of truncated *STMN2* RNA in the frontal cortex to discriminate between FTLD-TDP and non-FTLD-TDP cases (again separately for cognitively normal controls and PSP cases), we estimated area under the ROC curve (AUC); an AUC of 0.5 corresponds to predictive ability equal to that of chance, whereas an AUC of 1.0 indicates perfect predictive ability. We applied a Bonferroni correction to adjust for the five different statistical tests that were performed for the separate FTLD-TDP vs. controls and FTLD-TDP vs. PSP analyses, after which P values <0.01 were considered as statistically significant.

In secondary analysis, we compared truncated *STMN2* RNA between controls and PSP patients, and made pair-wise comparisons of full-length *STMN2* RNA and phosphorylated TDP-

43 (both on the base-2 logarithm scale) between controls, PSP cases, and FTLD-TDP cases using the aforementioned linear regression analyses.

In the subgroup of 238 diseased patients, associations of five different variables (phosphorylated TDP-43 levels, TDP-43 subtype, age at onset, survival after onset, and sex) with both truncated and full-length *STMN2* RNA levels were evaluated using single-variable and multivariable linear regression models, where truncated and full-length *STMN2* RNA in the frontal cortex were assessed on the base-2 logarithm scale, and multivariable models were adjusted for pre-defined potential confounding variables. The multivariable model assessing phosphorylated TDP-43 response in frontal cortex was adjusted for age at death, sex, and TDP-43 subtype. The multivariable model assessing TDP-43 subtype was adjusted for age at death and sex. The multivariable model assessing age at onset was adjusted for sex and TDP-43 subtype. The multivariable model assessing survival after onset was adjusted for age at onset, sex, and TDP-43 subtype. Finally, the multivariable model assessing sex was adjusted for age at death and TDP-43 subtype, and genotype. Regression coefficients and 95% CIs were estimated and are interpreted as the change in the mean level of truncated or full-length *STMN2* RNA in the frontal cortex (on the base-2 logarithm scale) corresponding to presence of the given characteristic (categorical variables) or a specified increase (continuous variables). Phosphorylated TDP-43 response in the frontal cortex was examined on the base-2 logarithm scale in linear regression analysis due to its skewed distribution. In order to adjust for the five different variables that were assessed for association with the separate truncated and full-length *STMN2* RNA outcomes, we utilized a Bonferroni correction for multiple testing, after which P-values <0.01 were considered as statistically significant.

All statistical tests were two-sided. All statistical analyses described above were performed using R Statistical Software (version 3.6.2; R Foundation for Statistical Computing, Vienna, Austria)

Statistical analyses for iPSC studies and qRT-PCR on RNA extracted from human tissues were performed using GraphPad Prism 8 software, where one-way analysis of variance (ANOVA) with Bonferroni's post-hoc test were performed for multiple group comparisons and unpaired Student *t*-tests were performed when only two groups were compared. The statistical test and number of independent experiments used for each analysis are indicated in each figure legend.

Study approval

The NYGC ALS Consortium samples presented in this work were acquired through various Institution Review Board protocols from member sites and the Target ALS post-mortem tissue core, transferred to NYGC in accordance with all applicable foreign, domestic, federal, state, and local laws and regulations for processing, sequencing, and analyses.

Post-mortem brain tissues from patients with FTLD-TDP, PSP and cognitively normal control individuals were obtained from the Mayo Clinic Florida Brain Bank. Diagnosis was independently ascertained by trained neurologists and neuropathologists upon neurological and pathological examinations, respectively. Written informed consent was given by all participants or authorized family members and all protocols were approved by the Mayo Clinic Institution Review Board and Ethics Committee. A summary of patient characteristics is described in Table 1. Sample size was determined by tissue availability and quality of RNA obtained from such samples, as described below.

Acknowledgments

We thank the Target ALS Postmortem Core for providing post-mortem brain samples and all patients and their families for their contribution to this study. See Supplemental Acknowledgments for consortium details.

Funding

This work was supported by the National Institutes of Health/National Institute of Neurological Disorder and Stroke [R35NS097273 (L.P.); P01NS084974 (L.P.); P01NS099114 (T.F.G., L.P.); R01NS088689 (L.P.); R21NS084528 (L.P.); R01AG037491 (K.A.J); P50AG016574 (R.C.P., B.F.B.); U01AG006786 (R.C.P.), U01AG045390 (B.F.B.)]; National Institute on Aging [R56AG055824 (J.H., T.R.)]; National Institute of Environmental Health Services [R01ES20395 (L.P.)]; Department of Defense [ALSRP AL130125 (L.P.)]; Mayo Clinic Foundation (L.P.); Amyotrophic Lateral Sclerosis Association (L.P., M.P.); Robert Packard Center for ALS Research at Johns Hopkins (L.P.); Target ALS (T.F.G., L.P.); Canadian Institute

of Health Research (S.P.); UK Medical Research Council (P.F., E.M.C.F.); Motor Neuron Disease Association (P.F.); Rosetrees Foundation (P.F., E.M.C.F.); National Institute of Health UCLH Biomedical Research Centre (P.F., A.L.B.); Alzheimer's Research UK (T.L.); UK Dementia Research Institute which receives its funding from DRI Ltd, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK (B.T.S, S.C); and supported, in part, by the Intramural Research Program, NINDS, NIH (M.W.). All NYGC ALS Consortium activities are supported by the ALS Association (15-LGCA-234) and the Tow Foundation.

Author contributions

M.P., J.H. and S.P. share first co-author position. The order was established based on their individual contributions and the importance of those contributions to this study. M.P. led human brain RNA and protein studies that made the basis of the manuscript. J.H. led the RNA-seq data analyses that complemented the RNA/protein studies and increased the impact of this manuscript. S.P. led iPSC studies and has been closely involved in the RNA/protein studies as well. M.P., J.H., S.P., T.F.G., P.F., and L.P. designed the study. M.P., S.P., A.-L.B., C.C., Y.S., M.Y., L.M.D., Y.C., K.J.-W., C.B., A.C., K.T., B.T.S., K.B., S.C., and E.M.C.F. generated iPSC lines, and processed iPSC and human tissue samples for RNA. S.E.H and M.W. contributed to the experimental design for CRISPRi knockdown of *TARDBP* in iPSC-derived neurons. Y.-C.W., I.H., D.C., D.K., N.P., S.F., D.F., and H.P. contributed with RNA-seq sample processing, de-identified clinical and genotype information acquisition, data distribution, and coordination of the research activity of NYGC ALS Consortium provided through the Target ALS Post-mortem Core. J.H., Y.-C.W., P.S., A.-L.B., M.S. and T.R. performed RNA-seq data analyses. M.P., J.M.K., J.S., A.E.T. performed NanoString assays and analyzed the data. S.P. and T.F.G. performed pTDP-43 immunoassay. M.P., M.G.H., M.R.S., A.-L.B., and M.S. performed statistical analyses. M.P., J.H., S.P., A.-L.B., T.F.G, P.F., and L.P wrote the manuscript. C.F.de C., M.D., S.K., J.N., T.L., NYGC ALS Consortium, B.O., M.V.B., R.R., N.R.G., B.F.B., D.S.K., R.C.P, K.A.J., and D.D provided human tissue samples, as well as pathological, genetic, and clinical information. All authors read and approved the manuscript.

Declaration of interests

R.C.P has the following disclosures: Roche, Inc.; Merck, Inc.; Biogen, Inc.; Eisai, Inc. consultant, Genentech, member of DSMB; GE Healthcare, educational talk.

Data and code availability

RNA-seq data generated through the NYGC ALS Consortium in this study can be accessed at GEO (GSE137810, GSE124439, GSE116622 and GSE153960). To request immediate access to new and ongoing data generated by the NYGC ALS Consortium for samples provided through the Target ALS post-mortem core, contact ALSData@nygenome.org. All RNA-seq data in the NYGC ALS Consortium are made immediately available to all members of the Consortium and with other Consortia with whom we have a reciprocal sharing arrangement.

FIGURE 1

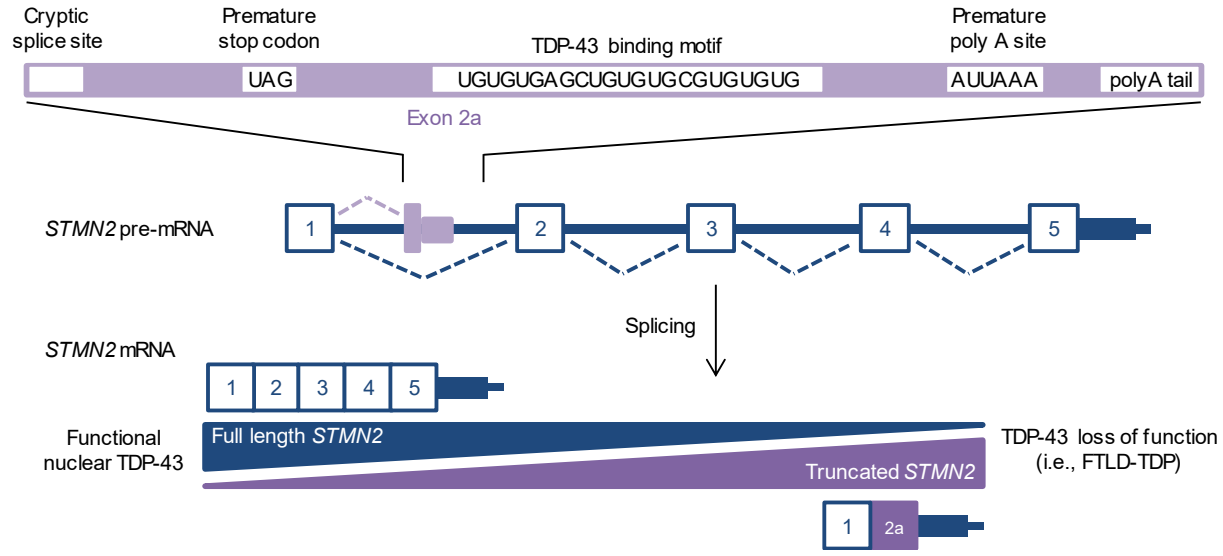

Figure 1. Truncated *STMN2* RNA is generated with loss of nuclear TDP-43. Schematic of TDP-43-regulated *STMN2* splicing. When TDP-43 is present and functional, it binds to GU-rich sequences in the first intron of the *STMN2* transcript and allows normal splicing of intron 1. Reduced TDP-43 binding to *STMN2* RNA, from TDP-43 aggregation or depletion from the nucleus, leads to mis-splicing of *STMN2* RNA, resulting in the inclusion of a novel exon encoded within intron 1 (termed exon 2a), and containing an alternative or cryptic polyadenylation site. The production of this alternative variant of *STMN2*, which lacks exons 2 through 5 (referred to as truncated *STMN2*), is at the expense of full-length *STMN2*, which is reduced upon TDP-43 downregulation.

FIGURE 2

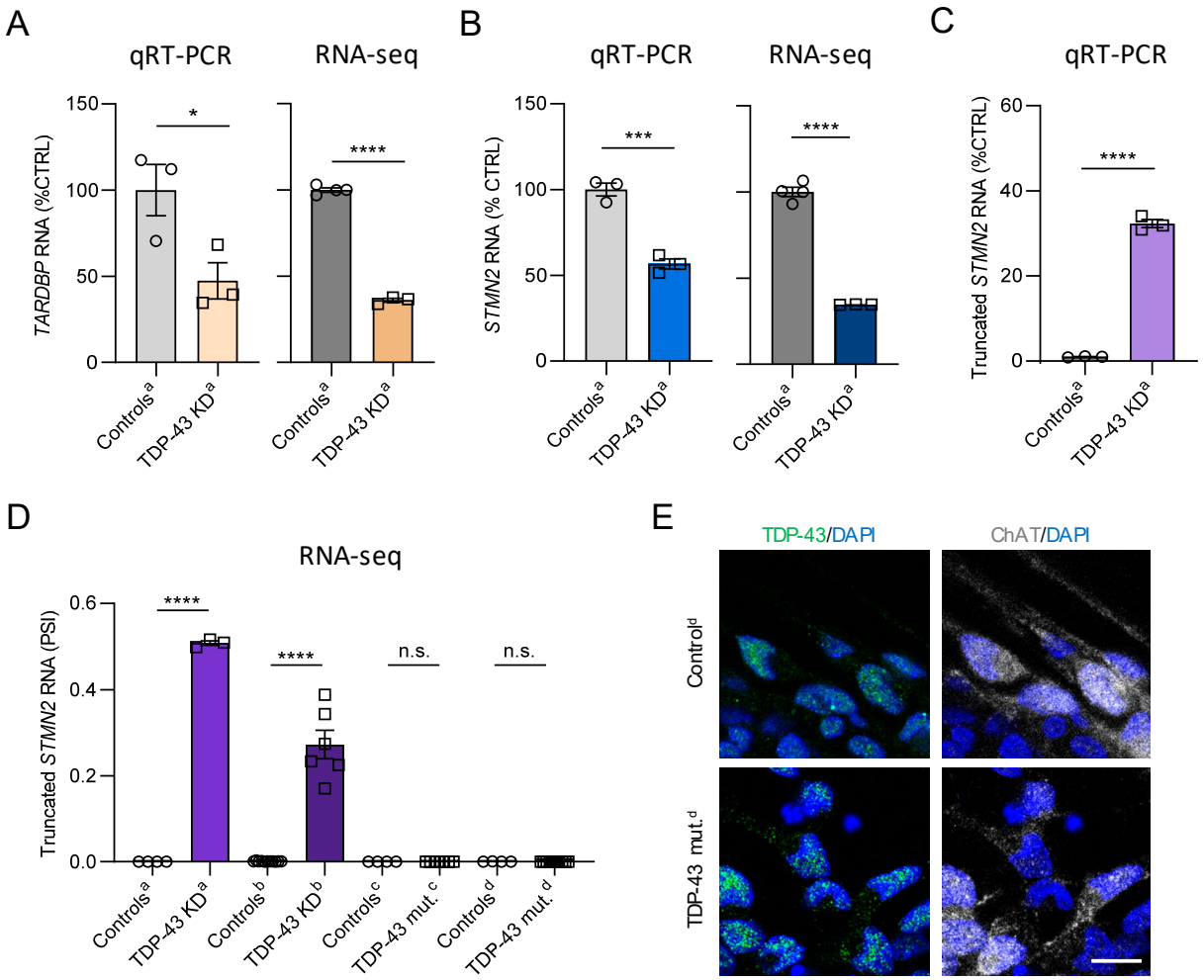

Figure 2. Truncated *STMN2* RNA is elevated in iPSC-derived neurons with reduced levels of TDP-43. (**A-C**) iPSC-derived cortical neurons constitutively expressing CRISPR inactivation machinery (dataset labeled "a") were transduced with lentivirus expressing a sgRNA against *TARDBP* (TDP-43 KD) or a control sgRNA (Controls) and subjected to analysis by qRT-PCR (N=3 TDP-43 KD, N=3 Controls) and/or RNA-seq (N=3 TDP-43 KD, N=4 Controls). Graphs demonstrate a decrease in *TARDBP* transcript (**A**), a reduction of the full-length *STMN2* transcript (**B**), and an increase in the truncated *STMN2* transcript (**C**) upon TDP-43 depletion. (**D**) Detection of *STMN2* cryptic exon 2a inclusion expressed as percent spliced in (PSI) by RNA-seq in various iPSC-derived neurons. iPSC-derived cortical neurons (dataset labeled "a") described above and performed in this study, iPSC-derived motor neurons treated with *TARDBP* siRNA (TDP-43 KD) or control siRNA (Controls) from Klim *et al.* (dataset labeled "b"; N=6 TDP-43 KD, N=11 Controls) (7), iPSC-derived motor neurons from patients with *TARDBP* mutations (TDP-43 mut.) or controls (Controls) from two independent groups in Edinburgh (dataset labeled "c"; N=7 TDP-43 KD, N=4 Controls) and Oxford (dataset labeled "d", N=10 TDP-43 KD, N=4 Controls) (73). Data is presented as mean ± SEM, and normalized to the control group in **A-B** (set to 100%). *P*-values result from unpaired Student *t* tests (**A-C**) or One-way ANOVA (**D**): *P<0.05, ***P<0.005, ****P<0.001, n.s.: not significant differences. (**E**) Representative images of iPSC-derived motor neurons from patients with *TARDBP* mutations

and controls from a group in Oxford (dataset labeled “d” in **Figure 1D**) show similar nuclear-cytoplasmic distribution of TDP-43 (green) at day 30 of differentiation (DAPI in blue, Choline acetyltransferase (ChAT) in white). Scale bar indicates 10 μm. All panels of Figure 2E are reshown in **Supplementary Figure 3A**. Additional data associated with this figure can also be found in **Supplementary Figure 3**.

FIGURE 3

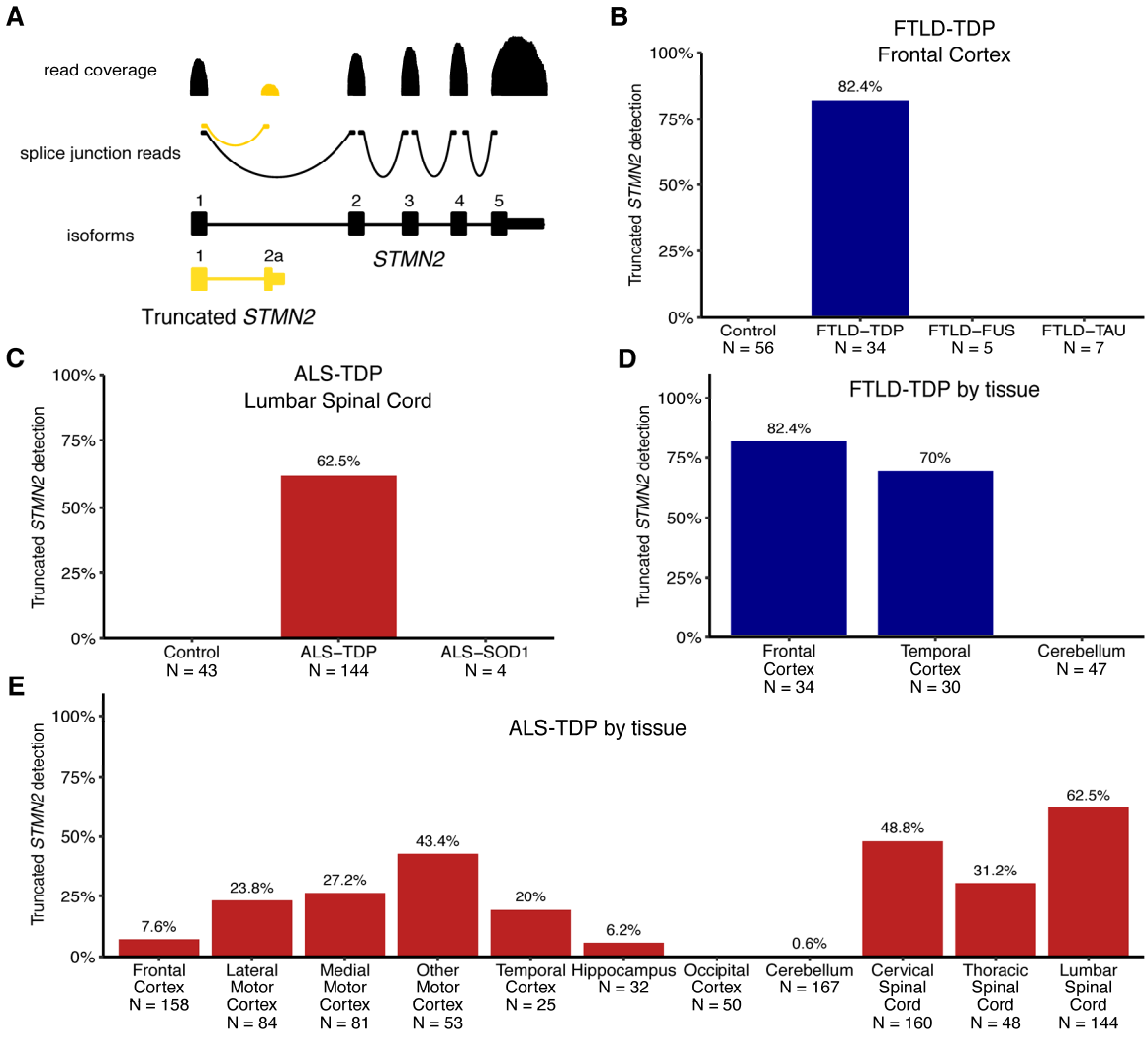

Figure 3. Truncated *STMN2* RNA is detected in bulk RNA-seq from FTLD/ALS tissues with TDP-43 pathology. (**A**) NYGC ALS Consortium RNA-seq dataset was analyzed for the presence of truncated *STMN2* transcripts, which were identified by RNA-seq reads spanning exon 1-exon 2a splice junction. (**B-E**) Bar graphs representing the proportion of individuals (%) with at least two reads spanning exon 1-exon 2a junction. N refers to the number of individuals with samples in indicated tissues/diseases. (**B**) Truncated *STMN2* was detected in the frontal cortex of FTLD-TDP cases but not in controls, FTLD-FUS or FTLD-tau. (**C**) In the lumbar spinal cord of ALS cases, truncated *STMN2* RNA is only detected in at least one tissue in ALS-TDP samples but not in ALS-SOD1 or controls. (**D**) Among FTLD-TDP tissues, truncated *STMN2* is only detected in tissues known to have TDP-43 pathology (frontal and temporal cortex), while (**E**) within ALS-TDP tissues, truncated *STMN2* RNA is seen in motor cortex and spinal cord tissues but not in regions spared of TDP-43 pathology.

FIGURE 4

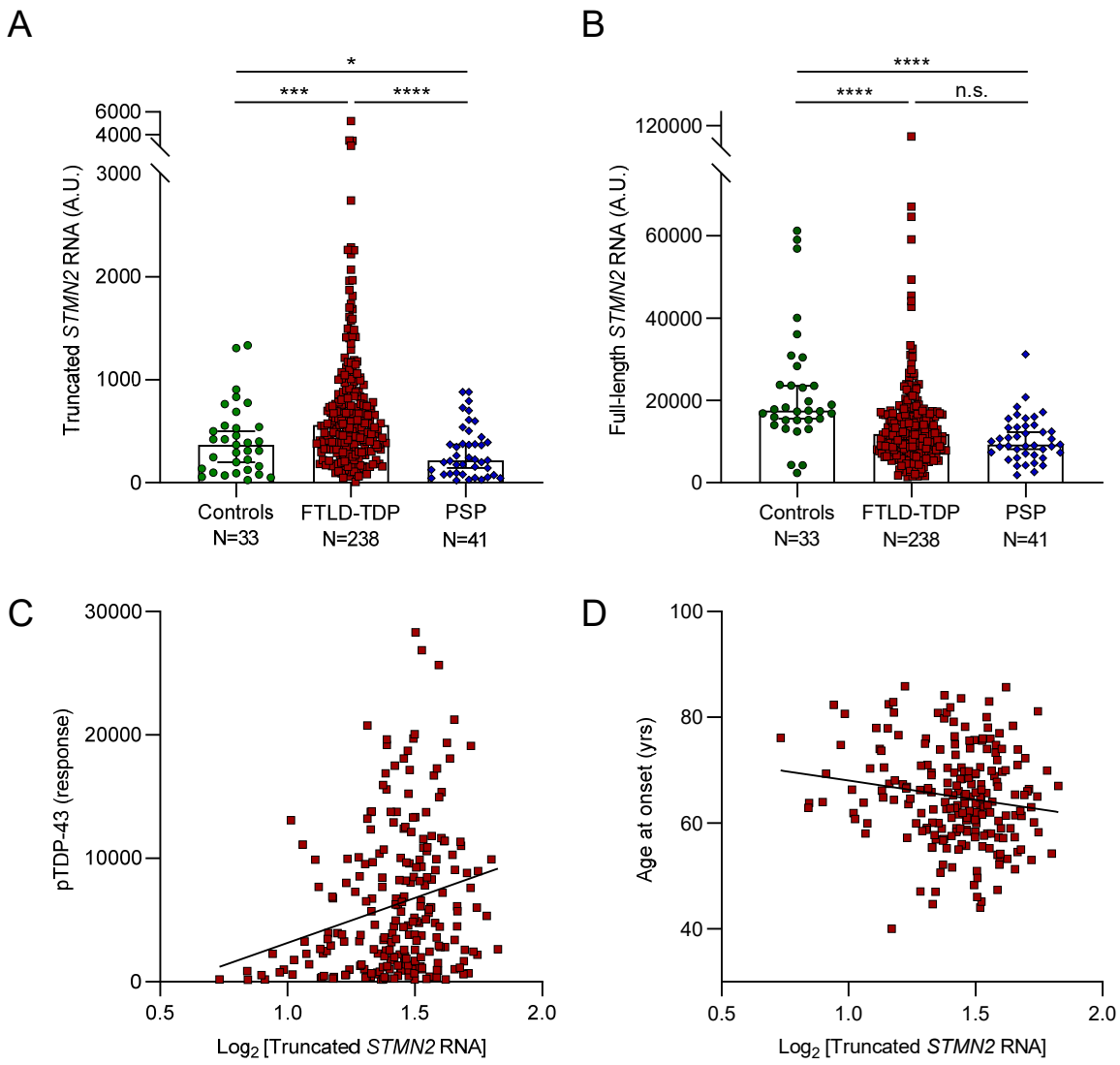

Figure 4. Truncated *STMN2* is elevated in FTLD-TDP and associates with higher pTDP-43 burden and earlier age of FTD onset. (A-B) The levels of truncated and full-length *STMN2* were measured from RNA extracted from frontal cortex of controls, FTLD-TDP and PSP cases, using NanoString Plexset platform. (**A**) A significant accumulation of truncated *STMN2* RNA was observed in FTLD-TDP cases but not in controls (see also **Supplementary Table 2**) or PSP cases (see also **Table 2**). (**B**) Full-length *STMN2* levels were significantly decreased in FTLD-TDP and PSP cases (see also **Supplementary Table 3**). Data is presented as median with 95% CI, *P*-values result from linear regression models that were adjusted for age at death, sex and RIN. *P<0.05, ***P<0.005, ****P<0.001, n.s.: not significant differences. (**C**-**D**) In our FTLD-TDP cohort (N=238), significant correlations were observed between truncated *STMN2* RNA levels and higher burden of pTDP-43 (**C**), and with an earlier age of disease onset (**D**). See **Table 3** for correlation coefficients.

Tables

Table 1. Patient characteristics of the post-mortem cohort.

Variable	Controls (N=33)	PSP (N=41)	All FTLD-TDP (N=238)	FTLD-TDP type A (N=117)	FTLD-TDP type B (N=66)	FTLD-TDP type C (N=43)	FTLD-TDP type D (N=2)
Age at death (years)	82 (55, 99)	69 (52, 86)	72 (49, 100)	76 (50, 100)	67 (49, 98)	74 (60, 84)	63 (61, 64)
Sex (female)	13 (39.4%)	20 (48.8%)	108 (45.4%)	54 (46.2%)	29 (43.9%)	20 (46.5%)	1 (50.0%)
Survival after onset (years)	NA	7.0 (4.0, 14.0)	7.0 (0.6, 25.0)	7.0 (1.0, 25.0)	4.0 (0.6, 20.0)	10.6 (4.0, 25.0)	8.5 (7.0, 10.0)
Age at onset (years)	NA	61.5 (46.0, 79.5)	64.5 (40.0, 85.8)	67.8 (40.0, 85.8)	61.4 (44.0, 85.7)	62.7 (46.0, 76.0)	54.1 (54.0, 54.2)
RIN	9.6 (7.1, 10.0)	9.5 (7.1, 10.0)	9.5 (7.0, 10.0)	9.4 (7.0, 10.0)	9.6 (7.0, 10.0)	9.5 (7.8, 10.0)	9.9 (9.8, 10.0)

The sample median (minimum, maximum) is given for continuous variables. Information was unavailable regarding TDP-43 subtype (All FTLD-TDP: N=10), age at death (All FTLD-TDP: N=1), pTDP43 burden in frontal cortex (All FTLD-TDP: N=1), survival after onset (PSP: N=4, All FTLD-TDP: N=20), and age at onset (PSP: N=4, All FTLD-TDP: N=20). Among the FTLD-TDP cases: 50 were *C9orf72* mutants (27 TDP-43 type A, 21 TDP-43 type B, two TDP-43 type C), 36 were *GRN* mutants (34 TDP-43 type A, one TDP-43 type B, one TDP-43 type C), and two were *VCP* mutants (TDP-43 type D). NA: not applicable.

Table 2. Comparisons of truncated *STMN2* RNA in frontal cortex between FTLD-TDP cases and PSP cases.

			Unadjusted analysis		Adjusting for age at death, sex, and RIN		
Group	**N**	**Median (minimum, maximum) levels of truncated *STMN2* RNA**	**Regression coefficient (95% CI)**	**P-value**	**Regression coefficient (95% CI)**	**P-value**	**AUC (95% CI) vs. PSP cases**
PSP cases	41	217 (22, 880)	0.00 (reference)	NA	0.00 (reference)	NA	NA
All FTLD-TDP	238	560 (8, 5210)	1.38 (0.96, 1.80)	<0.001	1.36 (0.95, 1.77)	<0.001	0.76 (0.69, 0.84)
FTLD-TDP type A	117	595 (8, 3495)	1.47 (0.98, 1.95)	<0.001	1.37 (0.87, 1.86)	<0.001	0.78 (0.70, 0.86)
FTLD-TDP type B	66	494 (62, 5210)	1.28 (0.77, 1.79)	<0.001	1.29 (0.79, 1.78)	<0.001	0.74 (0.65, 0.84)
FTLD-TDP type C	43	575 (106, 2071)	1.34 (0.79, 1.89)	<0.001	1.26 (0.73, 1.78)	<0.001	0.77 (0.66, 0.87)
FTLD-TDP type D	2	1035 (1008, 1062)	2.40 (0.25, 4.55)	0.029	3.07 (1.07, 5.07)	0.004	1.00 (1.00, 1.00)

CI=confidence interval; AUC=area under the ROC curve. Regression coefficients, 95% CIs, and P-values result from linear regression models where the levels of truncated *STMN2* RNA in the frontal cortex were considered on the base 2 logarithm scale. Regression coefficients are interpreted as the difference in mean levels of truncated *STMN2* RNA in the frontal cortex (on the base 2 logarithm scale) between the given group of diseased patients and PSP cases. P-values ≤ 0.010 are considered as statistically significant after applying a Bonferroni correction for multiple testing. NA: not applicable.

Table 3. Associations of truncated *STMN2* with pTDP-43 and clinical characteristics in FTLD-TDP frontal cortex.

Association with frontal cortex truncated *STMN2* RNA					
	Unadjusted analysis		Multivariable analysis		
Variable	**Regression coefficient (95% CI)**	**P-value**	**Regression coefficient (95% CI)**	**P-value**	**Multivariable model adjustments**
pTDP-43 (doubling)	0.20 (0.12, 0.28)	<0.001	0.19 (0.09, 0.28)	<0.001	Age at death, sex, TDP-43 subtype
TDP-43 subtype	Overall test of difference: P=0.48		Overall test of difference: P=0.33		Age at death and sex
FTLD-TDP type A	0.00 (reference)	NA	0.00 (reference)	NA	
FTLD-TDP type B	-0.19 (-0.57, 0.19)	0.33	-0.31 (-0.72, 0.10)	0.14	
FTLD-TDP type C	-0.13 (-0.56, 0.31)	0.57	-0.16 (-0.60, 0.27)	0.46	
FTLD-TDP type D	0.93 (-0.91, 2.77)	0.32	0.68 (-1.17, 2.53)	0.47	
Age at onset (10 year increase)	-0.23 (-0.41, -0.04)	0.015	-0.29 (-0.49, -0.09)	0.005	Sex and TDP-43 subtype
Survival after onset (5 year increase)	0.12 (-0.06, 0.30)	0.19	0.10 (-0.09, 0.29)	0.32	Age at onset, sex, and TDP-43 subtype
Sex (male)	-0.37 (-0.68, -0.06)	0.019	-0.35 (-0.67, -0.03)	0.034	Age at death and TDP-43 subtype

CI=confidence interval. Regression coefficients, 95% CIs, and P-values result from linear regression models, where frontal cortex truncated *STMN2* RNA was considered on the base 2 logarithm scale. Regression coefficients are interpreted as the change in mean frontal cortex truncated *STMN2* RNA (on the base 2 logarithm scale) corresponding to presence of the given characteristic (categorical variables) or the increase given in parenthesis (continuous variables). P-values ≤ 0.010 are considered as statistically significant after applying a Bonferroni correction for multiple testing. NA: not applicable.

References

1. Ling SC, Polymenidou M, and Cleveland DW. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron.* 2013;79(3):416-38.
2. Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling SC, Sun E, Wancewicz E, Mazur C, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci.* 2011;14(4):459-68.
3. Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, Konig J, Hortobagyi T, Nishimura AL, Zupunski V, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature neuroscience.* 2011;14(4):452-8.
4. Ling JP, Pletnikova O, Troncoso JC, and Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science.* 2015;349(6248):650-5.
5. Jeong YH, Ling JP, Lin SZ, Donde AN, Braunstein KE, Majounie E, Traynor BJ, LaClair KD, Lloyd TE, and Wong PC. Tdp-43 cryptic exons are highly variable between cell types. *Mol Neurodegener.* 2017;12(1):13.
6. Humphrey J, Emmett W, Fratta P, Isaacs AM, and Plagnol V. Quantitative analysis of cryptic splicing associated with TDP-43 depletion. *BMC Med Genomics.* 2017;10(1):38.
7. Klim JR, Williams LA, Limone F, Guerra San Juan I, Davis-Dusenbery BN, Mordes DA, Burberry A, Steinbaugh MJ, Gamage KK, Kirchner R, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci.* 2019;22(2):167-79.
8. Melamed Z, Lopez-Erauskin J, Baughn MW, Zhang O, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci.* 2019;22(2):180-90.
9. Shin JE, Geisler S, and DiAntonio A. Dynamic regulation of SCG10 in regenerating axons after injury. *Exp Neurol.* 2014;252(1-11.
10. Neumann M, and Mackenzie IRA. Review: Neuropathology of non-tau frontotemporal lobar degeneration. *Neuropathol Appl Neurobiol.* 2019;45(1):19-40.
11. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science.* 2008;319(5870):1668-72.
12. Borroni B, Bonvicini C, Alberici A, Buratti E, Agosti C, Archetti S, Papetti A, Stuani C, Di Luca M, Gennarelli M, et al. Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. *Hum Mutat.* 2009;30(11):E974-83.
13. Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K, Weng SM, Haass C, Kretzschmar HA, Edbauer D, et al. Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol.* 2013;126(6):859-79.
14. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science.* 2006;314(5796):130-3.
15. Sampathu DM, Neumann M, Kwong LK, Chou TT, Micsenyi M, Truax A, Bruce J, Grossman M, Trojanowski JQ, and Lee VM. Pathological heterogeneity of frontotemporal lobar degeneration with ubiquitin-positive inclusions delineated by ubiquitin immunohistochemistry and novel monoclonal antibodies. *Am J Pathol.* 2006;169(4):1343-52.

16. Davidson Y, Kelley T, Mackenzie IR, Pickering-Brown S, Du Plessis D, Neary D, Snowden JS, and Mann DM. Ubiquitinated pathological lesions in frontotemporal lobar degeneration contain the TAR DNA-binding protein, TDP-43. *Acta Neuropathol (Berl).* 2007;113(5):521-33.
17. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, 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.* 2006;351(3):602-11.
18. Hatanpaa KJ, Bigio EH, Cairns NJ, Womack KB, Weintraub S, Morris JC, Foong C, Xiao G, Hladik C, Mantanona TY, et al. TAR DNA-binding protein 43 immunohistochemistry reveals extensive neuritic pathology in FTLD-U: a midwest-southwest consortium for FTLD study. *J Neuropathol Exp Neurol.* 2008;67(4):271-9.
19. Josephs KA, Stroh A, Dugger B, and Dickson DW. Evaluation of subcortical pathology and clinical correlations in FTLD-U subtypes. *Acta Neuropathol.* 2009;118(3):349-58.
20. Mackenzie IR, Neumann M, Baborie A, Sampathu DM, Du Plessis D, Jaros E, Perry RH, Trojanowski JQ, Mann DM, and Lee VM. A harmonized classification system for FTLD-TDP pathology. *Acta Neuropathol.* 2011;122(1):111-3.
21. Lee EB, Porta S, Michael Baer G, Xu Y, Suh E, Kwong LK, Elman L, Grossman M, Lee VM, Irwin DJ, et al. Expansion of the classification of FTLD-TDP: distinct pathology associated with rapidly progressive frontotemporal degeneration. *Acta Neuropathol.* 2017;134(1):65-78.
22. Mackenzie IR, Baborie A, Pickering-Brown S, Du Plessis D, Jaros E, Perry RH, Neary D, Snowden JS, and Mann DM. Heterogeneity of ubiquitin pathology in frontotemporal lobar degeneration: classification and relation to clinical phenotype. *Acta Neuropathol.* 2006;112(5):539-49.
23. Tan RH, Shepherd CE, Kril JJ, McCann H, McGeachie A, McGinley C, Affleck A, and Halliday GM. Classification of FTLD-TDP cases into pathological subtypes using antibodies against phosphorylated and non-phosphorylated TDP43. *Acta Neuropathol Commun.* 2013;1(33.
24. Josephs KA, Hodges JR, Snowden JS, Mackenzie IR, Neumann M, Mann DM, and Dickson DW. Neuropathological background of phenotypical variability in frontotemporal dementia. *Acta Neuropathol.* 2011;122(2):137-53.
25. Josephs KA, Ahmed Z, Katsuse O, Parisi JF, Boeve BF, Knopman DS, Petersen RC, Davies P, Duara R, Graff-Radford NR, et al. Neuropathologic features of frontotemporal lobar degeneration with ubiquitin-positive inclusions with progranulin gene (PGRN) mutations. *J Neuropathol Exp Neurol.* 2007;66(2):142-51.
26. Mackenzie IR, Baker M, Pickering-Brown S, Hsiung GY, Lindholm C, Dwosh E, Gass J, Cannon A, Rademakers R, Hutton M, et al. The neuropathology of frontotemporal lobar degeneration caused by mutations in the progranulin gene. *Brain.* 2006;129(Pt 11):3081-90.
27. Feneberg E, Gray E, Ansorge O, Talbot K, and Turner MR. Towards a TDP-43-Based Biomarker for ALS and FTLD. *Mol Neurobiol.* 2018;55(10):7789-801.
28. Nana AL, Sidhu M, Gaus SE, Hwang JL, Li L, Park Y, Kim EJ, Pasquini L, Allen IE, Rankin KP, et al. Neurons selectively targeted in frontotemporal dementia reveal early stage TDP-43 pathobiology. *Acta Neuropathol.* 2019;137(1):27-46.

29. Koga S, Kouri N, Walton RL, Ebbert MTW, Josephs KA, Litvan I, Graff-Radford N, Ahlskog JE, Uitti RJ, van Gerpen JA, et al. Corticobasal degeneration with TDP-43 pathology presenting with progressive supranuclear palsy syndrome: a distinct clinicopathologic subtype. *Acta Neuropathol.* 2018;136(3):389-404.
30. Tian R, Gachechiladze MA, Ludwig CH, Laurie MT, Hong JY, Nathaniel D, Prabhu AV, Fernandopulle MS, Patel R, Abshari M, et al. CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron.* 2019.
31. Fratta P, Sivakumar P, Humphrey J, Lo K, Ricketts T, Oliveira H, Brito-Armas JM, Kalmar B, Ule A, Yu Y, et al. Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J.* 2018;37(11).
32. Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, Seelman A, Stauffer JE, Jafar-Nejad P, Drenner K, Schulte D, et al. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron.* 2016;90(3):535-50.
33. Prudencio M, Gonzales PK, Cook CN, Gendron TF, Daughrity LM, Song Y, Ebbert MTW, van Blitterswijk M, Zhang YJ, Jansen-West K, et al. Repetitive element transcripts are elevated in the brain of C9orf72 ALS/FTLD patients. *Hum Mol Genet.* 2017;26(17):3421-31.
34. Schludi MH, Becker L, Garrett L, Gendron TF, Zhou Q, Schreiber F, Popper B, Dimou L, Strom TM, Winkelmann J, et al. Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol.* 2017;134(2):241-54.
35. Bang J, Spina S, and Miller BL. Frontotemporal dementia. *Lancet.* 2015;386(10004):1672-82.
36. Wang Q, Zhang Y, Wang M, Song WM, Shen Q, McKenzie A, Choi I, Zhou X, Pan PY, Yue Z, et al. The landscape of multiscale transcriptomic networks and key regulators in Parkinson's disease. *Nat Commun.* 2019;10(1):5234.
37. Amador-Ortiz C, Lin WL, Ahmed Z, Personett D, Davies P, Duara R, Graff-Radford NR, Hutton ML, and Dickson DW. TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Ann Neurol.* 2007;61(5):435-45.
38. Arai T, Mackenzie IR, Hasegawa M, Nonoka T, Niizato K, Tsuchiya K, Iritani S, Onaya M, and Akiyama H. Phosphorylated TDP-43 in Alzheimer's disease and dementia with Lewy bodies. *Acta Neuropathol.* 2009;117(2):125-36.
39. Higashi S, Iseki E, Yamamoto R, Minegishi M, Hino H, Fujisawa K, Togo T, Katsuse O, Uchikado H, Furukawa Y, et al. Concurrence of TDP-43, tau and alpha-synuclein pathology in brains of Alzheimer's disease and dementia with Lewy bodies. *Brain Res.* 2007;1184(284-94.
40. Hu WT, Josephs KA, Knopman DS, Boeve BF, Dickson DW, Petersen RC, and Parisi JE. Temporal lobar predominance of TDP-43 neuronal cytoplasmic inclusions in Alzheimer disease. *Acta Neuropathol.* 2008;116(2):215-20.
41. Josephs KA, Whitwell JL, Knopman DS, Hu WT, Stroh DA, Baker M, Rademakers R, Boeve BF, Parisi JE, Smith GE, et al. Abnormal TDP-43 immunoreactivity in AD modifies clinicopathologic and radiologic phenotype. *Neurology.* 2008;70(19 Pt 2):1850-7.

42. Kadokura A, Yamazaki T, Kakuda S, Makioka K, Lemere CA, Fujita Y, Takatama M, and Okamoto K. Phosphorylation-dependent TDP-43 antibody detects intraneuronal dot-like structures showing morphological characters of granulovacuolar degeneration. *Neurosci Lett.* 2009;463(1):87-92.
43. Kadokura A, Yamazaki T, Lemere CA, Takatama M, and Okamoto K. Regional distribution of TDP-43 inclusions in Alzheimer disease (AD) brains: their relation to AD common pathology. *Neuropathology.* 2009;29(5):566-73.
44. Uryu K, Nakashima-Yasuda H, Forman MS, Kwong LK, Clark CM, Grossman M, Miller BL, Kretzschmar HA, Lee VM, Trojanowski JQ, et al. Concomitant TAR-DNA-binding protein 43 pathology is present in Alzheimer disease and corticobasal degeneration but not in other tauopathies. *J Neuropathol Exp Neurol.* 2008;67(6):555-64.
45. Josephs KA, Dickson DW, Tosakulwong N, Weigand SD, Murray ME, Petrucelli L, Liesinger AM, Senjem ML, Spychalla AJ, Knopman DS, et al. Rates of hippocampal atrophy and presence of post-mortem TDP-43 in patients with Alzheimer's disease: a longitudinal retrospective study. *Lancet Neurol.* 2017;16(11):917-24.
46. Josephs KA, Whitwell JL, Tosakulwong N, Weigand SD, Murray ME, Liesinger AM, Petrucelli L, Senjem ML, Ivnik RJ, Parisi JE, et al. TAR DNA-binding protein 43 and pathological subtype of Alzheimer's disease impact clinical features. *Ann Neurol.* 2015;78(5):697-709.
47. Chen S, Zhou Y, Chen Y, and Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884-i90.
48. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.
49. Liao Y, Smyth GK, and Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923-30.
50. Robinson MD, McCarthy DJ, and Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139-40.
51. Katz Y, Wang ET, Airoldi EM, and Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods.* 2010;7(12):1009-15.
52. Tam OH, Rozhkov NV, Shaw R, Kim D, Hubbard I, Fennessey S, Propp N, Consortium NA, Fagegaltier D, Harris BT, et al. Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *Cell Rep.* 2019;29(5):1164-77 e5.
53. Wang Y-C, Castellanos R, Pandya C, and Giles Z. RAPiD: An Agile and Dependable RNA-Seq Framework. *The 65th Annual Meeting of The American Society of Human Genetics, Baltimore, MD.* 2015.
54. Nextflow Data-driven computational pipelines.
55. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.
56. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 2019;47(D1):D766-D73.
57. Li B, and Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12(323.

58. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-9.
59. Picard Toolkit. *Broad Institute, GitHub Repository.* 2019.
60. Ewels P, Magnusson M, Lundin S, and Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016;32(19):3047-8.
61. Team RC. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* 2019.
62. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. 2017.
63. Kassambara A. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2.2. 2019.
64. Pedersen TL. patchwork: The Composer of Plots. R package version 1.0.0. 2019.
65. Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, Kasela S, Kim-Hellmuth S, Liang Y, Oliva M, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *bioRxiv.* 2019.
66. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, and Ragg T. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol.* 2006;7(3.
67. Jaffe AE, Tao R, Norris AL, Kealhofer M, Nellore A, Shin JH, Kim D, Jia Y, Hyde TM, Kleinman JE, et al. qSVA framework for RNA quality correction in differential expression analysis. *Proc Natl Acad Sci U S A.* 2017;114(27):7130-5.
68. Koster J, and Rahmann S. Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics.* 2012;28(19):2520-2.
69. STMN2 splicing pipeline. https://github.com/RajLabMSSM/STMN2-splicing. Accessed April 10, 2020.
70. Feng Y-Y, Ramu A, Cotto KC, Skidmore ZL, Kunisaki J, Conrad DF, Lin Y, Chapman WC, Uppaluri R, Govindan R, et al. RegTools: Integrated analysis of genomic and transcriptomic data for discovery of splicing variants in cancer. *bioRxiv.* 2018.
71. Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, and Pritchard JK. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet.* 2018;50(1):151-8.
72. Jackson JL, Finch NA, Baker MC, Kachergus JM, DeJesus-Hernandez M, Pereira K, Christopher E, Prudencio M, Heckman MG, Thompson EA, et al. Elevated methylation levels, reduced expression levels, and frequent contractions in a clinical cohort of C9orf72 expansion carriers. *Mol Neurodegener.* 2020;15(1):7.
73. Dafinca R, Barbagallo P, Farrimond L, Candalija A, Scaber J, Ababneh NA, Sathyaprakash C, Vowles J, Cowley SA, and Talbot K. Impairment of Mitochondrial Calcium Buffering Links Mutations in C9ORF72 and TARDBP in iPS-Derived Motor Neurons from Patients with ALS/FTD. *Stem Cell Reports.* 2020;14(5):892-908.