Common ALS/FTD risk variants in *UNC13A* exacerbate its cryptic splicing and loss upon TDP-43 mislocalization

Anna-Leigh Brown
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
Oscar Wilkins
The Francis Crick Institute
Matthew Keuss
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
Sarah Hill
NIH https://orcid.org/0000-0002-8499-7455
Matteo Zanovello
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
Weaverly Lee
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
Flora Lee
The Francis Crick Institute https://orcid.org/0000-0002-3216-9754
Laura Masino
The Francis Crick Institute
Yue Qi
National Institute of Neurological Disorders and Stroke, NIH
Sam Byrce-Smith
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
Alexander Bampton
Queen Square Brain Bank, UCL Queen Square Institute of Neurology, University College London
Ariana Gatt
Queen Square Brain Bank, UCL Queen Square Institute of Neurology, University College London
Hemali Phatnani
New York Genome Center
Giampietro Schiavo
UCL-Institute of Neurology London UK https://orcid.org/0000-0002-4319-8745
Elizabeth Fisher
UCL https://orcid.org/0000-0003-2850-9936
Towfique Raj

Icahn School of Medicine at Mount Sinai
Maria Secrier
University College London https://orcid.org/0000-0003-2758-1741
Tammaryn Lashley
University College London
Jernej Ule
The Francis Crick Institute https://orcid.org/0000-0002-2452-4277
Emanuele Buratti
ICGEB https://orcid.org/0000-0002-1356-9074
Jack Humphrey
Icahn School of Medicine at Mount Sinai
Michael Ward
National Institute of Neurological Disorders and Stroke, NIH
Pietro Fratta (**p.fratta@ucl.ac.uk**)
Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology
https://orcid.org/0000-0002-8762-8188

Biological Sciences - Article

Keywords: UNC13A gene, TDP-43 mislocalization

Posted Date: April 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-386600/v1

Version of Record: A version of this preprint was published at Nature on February 23rd, 2022. See the published version at https://doi.org/10.1038/s41586-022-04436-3.

Abstract

Variants within the *UNC13A* gene have long been known to increase risk of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), two related neurodegenerative diseases defined by mislocalization of the RNA-binding protein TDP-43. Here, we show that TDP-43 depletion induces robust inclusion of a cryptic exon (CE) within *UNC13A*, a critical synaptic gene, resulting in nonsense-mediated decay and protein loss. Strikingly, two common polymorphisms strongly associated with ALS/FTD risk directly alter TDP-43 binding within the CE or downstream intron, increasing CE inclusion in cultured cells and in patient brains. Our findings, which are the first to demonstrate a genetic link specifically between loss of TDP-43 nuclear function and disease, reveal both the mechanism by which *UNC13A* variants exacerbate the effects of decreased nuclear TDP-43 function, and provide a promising therapeutic target for TDP-43 proteinopathies.

One-sentence Summary

Shared ALS/FTD risk variants increase the sensitivity of a cryptic exon in the synaptic gene *UNC13A* to TDP-43 depletion.

Introduction

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are devastating adult-onset neurodegenerative disorders with shared genetic causes and common pathological aggregates[1–3]. Genome-wide association studies (GWASs) have repeatedly demonstrated a shared risk locus between ALS and FTD within the crucial synaptic gene *UNC13A*, although the mechanism underlying this association has remained elusive[4].

ALS and FTD are pathologically defined by cytoplasmic aggregation and nuclear depletion of TAR DNA-binding protein 43 (TDP-43) in the vast majority (> 97%) of ALS cases and in 45% of FTD cases (FTLD-TDP)[5]. TDP-43, an RNA-binding protein (RBP), primarily resides in the nucleus and plays key regulatory roles in RNA metabolism, including acting as a splicing repressor. Upon TDP-43 nuclear depletion – an early pathological feature in ALS/FTLD-TDP – non-conserved intronic sequences are de-repressed and erroneously included in mature RNAs. These events are referred to as cryptic exons (CEs) and can lead to premature stop-codons/polyadenylation and transcript degradation[6, 7]. Recently, TDP-43 loss was found to induce a CE in the *Stathmin 2* (*STMN2*) transcript, which can serve as a functional readout for TDP-43 proteinopathy, as it appears selectively in affected patient tissue and its level correlates with TDP-43 phosphorylation[8–10].

In this study, we report a novel CE in *UNC13A* which promotes nonsense-mediated decay, and is present at remarkably high levels in patient neurons. Strikingly, we find that ALS/FTD risk-associated SNPs within *UNC13A* promote increased inclusion of this CE. We thus elucidate the molecular mechanism behind one

of the top GWAS hits for ALS/FTD, and provide a promising new therapeutic target for TDP-43 proteinopathies.

Results

TDP-43 knockdown leads to inclusion of a cryptic exon in UNC13A

To discover novel CEs induced by TDP-43 depletion, we performed RNA-seq on human induced pluripotent stem cell (iPSC)-derived cortical-like i^{3}Neurons in which we reduced TDP-43 expression through CRISPR inhibition (CRISPRi)[10–13]. We identified 179 CEs, including several previously reported, such as *AGRN*, *PFKP* and *STMN2*[6–9] (Fig. 1A; data S1) (Fig. 1B; data S2). Interestingly, we observed robust mis-splicing in two members of the UNC13 synaptic protein family, *UNC13A* and *UNC13B* (Fig. 1C-F). Notably, *UNC13A* polymorphisms modify both disease risk and progression in ALS and FTLD-TDP [4, 14–21] pointing towards a potential functional relationship between TDP-43, *UNC13A*, and disease risk.

Inspection of the *UNC13A* gene revealed a previously unreported CE after TDP-43 knockdown (KD), with both a shorter and longer form, between exons 20 and 21 (Fig. 1C), and increased IR between exons 31 and 32 (fig. S1B). One ALS/FTLD-TDP risk SNP – *rs12973192*[17] – lies 16 bp inside the CE (henceforth referred to as the CE SNP). Another SNP – *rs12608932*[4] – is located 534 bp downstream of the donor splice site of the CE inside the same intron (henceforth referred to as the intronic SNP) (Fig. 1D). While there are five polymorphisms associated with ALS risk along *UNC13A*, they are all in high linkage disequilibrium (LD) in European populations with both the CE and intronic SNPs, and are present in 35% of individuals (Fig. 1G)[17]. The close proximity of the disease-associated SNPs to the *UNC13A* CE suggests that the SNPs may influence *UNC13A* splicing. In *UNC13B*, TDP-43 KD led to the inclusion of an annotated frame-shift-inducing exon between exons 10 and 11, henceforth referred to as the *UNC13B* frameshift exon (fsE), and increased intron retention (IR) between exon 21 and 22 (Fig. 1E,F; fig. S1A).

In support of a direct role for TDP-43 regulation of *UNC13A* and *UNC13B*, we found multiple TDP-43 binding peaks both downstream and within the body of the *UNC13A* CE (Fig. 1D) and IR (fig. S1B)[22], and *UNC13A* CE inclusion negatively correlated with *TARDBP* RNA levels (rho = -0.43, p = 0.077, Fig. 1H). Additionally, TDP-43 binding peaks were present near both splice events in *UNC13B* (Fig. 1F; fig. S1A)[22]. We also detected these splicing changes in RNA-seq from TDP-43 depleted SH-SY5Y and SK-N-DZ neuronal lines, as well as in publicly available iPSC-derived motor neurons (MNs)[8] and SK-N-DZ datasets[23](Fig. 1I-L; fig. S1C), and validated them by PCR in SH-SY5Y and SK-N-DZ cell lines (fig. S1D,E).

UNC13A and UNC13B RNA and protein are downregulated by TDP-43 knockdown

Next, we examined whether incorrect splicing of *UNC13A* and *UNC13B* affected transcript levels in neurons and neuron-like cells. TDP-43 KD significantly reduced *UNC13A* RNA abundance in the three cell types with the highest levels of cryptic splicing (FDR < 0.1; Fig. 2A, Fig. 1I). Likewise, *UNC13B* RNA was significantly downregulated in four datasets (FDR < 0.1) (Fig. 2B). We confirmed these results by qPCR in SH-SY5Y and SK-N-DZ cell lines (fig. S2A). The number of ribosome footprints aligning to *UNC13A* and

UNC13B was reduced after TDP-43 KD (Fig. 2C; fig. S2B, data S3). TDP-43 KD also decreased expression of UNC13A and UNC13B at the protein level, as assessed by quantitative proteomics with liquid chromatography tandem-mass spectrometry and western blot (Fig. 2D,E). These data suggest that the mis-splicing in *UNC13A* and *UNC13B* after TDP-43 KD reduces their transcript and protein abundance in neurons.

The *UNC13A* CE contains a premature termination codon (PTC) and is thus predicted to promote nonsense-mediated decay (NMD). Cycloheximide (CHX) treatment, which stalls translation and impairs NMD, increased CE inclusion in *UNC13A* after TDP-43 KD. Conversely, CHX did not alter levels of the aberrant *STMN2* transcript, which is not predicted to undergo NMD (Fig. 2F). Taken together, our data suggests that TDP-43 is critical for maintaining normal expression of the presynaptic proteins *UNC13A* and *UNC13B* by ensuring their correct pre-mRNA splicing.

UNC13A cryptic exon is highly expressed in TDP-43-depleted patient neurons

To explore whether the *UNC13A* CE could be detected in patient tissues affected by TDP-43 pathology, we first analysed RNA-seq from neuronal nuclei sorted from frontal cortices of ALS/FTLD patients[24]. We compared levels of *UNC13A* CE to levels of a CE in *STMN2* known to be regulated by TDP-43. Both *STMN2* and *UNC13A* CEs were exclusive to TDP-43-depleted nuclei, and, strikingly, in some cases the *UNC13A* CE percent spliced in (PSI) reached 100% (Fig. 3A). This suggests that in patients there will be a significant loss of UNC13A expression within the subpopulation of neurons with TDP-43 pathology.

Next, we quantified *UNC13A* CE inclusion in bulk RNA-seq from the NYGC ALS Consortium, a dataset containing 1,349 brain and spinal cord tissues from a total of 377 ALS, FTLD, and control individuals. The *UNC13A* CE was detected exclusively in FTLD-TDP and ALS-TDP cases (89% and 38% respectively), with no detection in ALS-non-TDP (*SOD1* and *FUS* mutations), FTLD-non-TDP (FTLD-TAU and FTLD-FUS), or control cases. The lower detection rate in ALS versus FTLD is likely due to the lower expression of *UNC13A* in the spinal cord (fig. S3A). Thus, pathological *UNC13A* CEs occur *in vivo* and are specific to neurodegenerative disease subtypes in which mislocalization and nuclear depletion of TDP-43 occurs.

UNC13A CE expression mirrored the known tissue distribution of TDP-43 aggregation and nuclear clearance[25]: it was specific to ALS-TDP spinal cord and motor cortex, as well as FTLD-TDP frontal and temporal cortices, but absent from the cerebellum in disease and control states (Fig. 3B). Despite the CE PSI being diluted by both the presence of unaffected cells and NMD in bulk RNA-seq, we were still able to detect CE above 20% in some samples. Furthermore, although, unlike the STMN2 CE, the *UNC13A* CE induces NMD, it was detected at similar levels to *STMN2* CE in cortical regions, whilst *STMN2* CE was more abundant in the spinal cord (Fig. 3C). We next investigated whether *UNC13A* CEs could be visualised by *in situ* hybridisation (ISH) in FTLD patient brains. Using a probe targeting the *UNC13A* CE on frozen frontal cortex tissue, we detected staining significantly above background in 4 out of 5 tested FTLD-TDP cases, but in none of the FTLD-Tau (n = 3) or control (n = 5) cases (Fig. 3D; fig. S3B).

To assess if *UNC13A* CE levels in bulk tissue was related to the level of TDP-43 proteinopathy, we used *STMN2* CE PSI as a proxy, as it correlates with the burden of phosphorylated TDP-43 in patient samples[10]. As expected, across the NYGC ALS Consortium samples we observed a significant positive correlation between the level of *STMN2* CE PSI and *UNC13A* CE PSI (rho = 0.55, p = 3.0e-4) (Fig. 3E). Combined, our analysis reveals a strong relationship between TDP-43 pathology and *UNC13A* CE levels, supporting a model for direct regulation of *UNC13A* mRNA splicing by TDP-43 in patients.

rs12973192(G) and rs12608932(C) combine to promote cryptic splicing

To test whether the ALS/FTD *UNC13A* risk SNPs promote cryptic splicing, which could explain their link to disease, we assessed *UNC13A* CE levels across different genotypes, and found significantly increased levels in cases homozygous for CE *rs12973192(G)* and intronic *rs12608932(C)* SNPs (fig. S4A-B). To ensure that this was not simply due to more severe TDP-43 pathology in these samples, we normalised by the level of *STMN2* cryptic splicing, and again found a significantly increased level of the *UNC13A* CE in cases with homozygous risk variants (Wilcoxon test, $p < 0.001$) (Fig. 4A; fig. S4C,D). Next, we performed targeted RNA-seq on *UNC13A* CE from temporal cortices of ten heterozygous risk allele cases and four controls. We detected significant biases towards reads containing the risk allele ($p < 0.05$, single-tailed binomial test) in six samples, with a seventh sample approaching significance (Fig. 4B), suggesting that the two ALS/FTLD-linked variants promote cryptic splicing *in vivo*.

To specifically examine whether the CE or the intronic SNP of *UNC13A* promote CE splicing, we generated four variants of minigenes containing *UNC13A* exon 20, intron 20, and exon 21, featuring both risk alleles (2R), both non-risk alleles (2H), the risk allele within the CE (*rs12973192*) (RE), or the risk allele in the intron (*rs12608932*) (RI) (Fig. 4C). We then expressed these minigenes in SH-SY5Y cells with doxycycline-inducible TDP-43 knockdown. We found that both the CE SNP and, to a lesser extent, intronic SNP independently promoted CE inclusion, with the greatest overall levels detected for the 2R minigene (Fig. 4D,E).

To explore how these two SNPs might act to enhance CE splicing, we analyzed a dataset of *in vitro* RNA heptamer/RBP binding enrichments, and examined the effect of the SNPs on relative RBP enrichment[26]. Strikingly, when investigating which RBPs were most impacted in their RNA binding enrichment by the CE-risk SNP, TDP-43 had the third largest decrease of any RBP, with only two non-human RBPs showing a larger decrease (Fig. 4F,G; fig. S4E,F). To test whether the CE SNP directly inhibited *in vitro* TDP-43 binding, we performed isothermal titration calorimetry using recombinant TDP-43 and 14-nt RNAs. As predicted, we observed an increased Kd for RNA containing the CE risk SNP (Fig. 4H; fig. S4G,H; Data S4). Together these data predict that the *UNC13A* CE SNP may directly inhibit TDP-43 binding.

To directly study the impact of the SNPs on TDP-43 binding to *UNC13A* pre-mRNA, we performed TDP-43 iCLIP with cells expressing either the 2R or 2H minigene. We observed a striking enrichment of crosslinks within the ~ 800nt UG-rich region containing both SNPs in intron 20 (Fig. 4I). When comparing the 2R with the 2H minigene, the peaks with the largest fractional changes were in close proximity of each SNP; similarly, we detected a 21% decrease in total TDP-43 crosslinks centred around the CE SNP and a 73%

increase upstream of the intronic SNP (Fig. 4I, J, fig. S4I; 50 nucleotide windows). These data demonstrate that these two disease-risk SNPs distort the pattern of TDP-43/RNA interactions, decreasing TDP-43 binding near the CE donor splice site, thus exacerbating *UNC13A* CE inclusion upon nuclear TDP-43 depletion.

Discussion

Our results support a model wherein TDP-43 nuclear depletion and the intronic and CE SNPs in *UNC13A* synergistically reduce expression of *UNC13A*, a gene that is critical for normal neuronal function. In this model, when nuclear TDP-43 levels are normal in healthy individuals, TDP-43 efficiently binds to *UNC13A* pre-mRNA and prevents CE splicing, regardless of *UNC13A* SNPs. Conversely, severe nuclear depletion of TDP-43 in end-stage disease induces CE inclusion in all cases. However, the common intronic and CE SNPs in *UNC13A* alter TDP-43 binding to *UNC13A* pre-mRNA and may make *UNC13A* CE more sensitive to partial TDP-43 loss that occurs early in degenerating neurons, explaining their associated risk effect. Strikingly, we found that both risk alleles for these SNPs independently and additively promoted cryptic splicing *in vitro.* Intriguingly, when the two variants are not co-inherited, as seen in East Asian individuals with ALS, an attenuated effect is observed[20]. A similar phenomenon wherein SNP pairs both contribute to risk has been widely studied at the APOE locus in Alzheimer's disease[27].

UNC13-family proteins are highly conserved across metazoans and are essential for calcium-triggered synaptic vesicle release[28]. In mice, double knockout of *UNC13A* and *UNC13B* inhibits both excitatory and inhibitory synaptic transmission in hippocampal neurons and greatly impairs transmission at neuromuscular junctions[29,30]. In TDP-43-negative neuronal nuclei derived from patients, the *UNC13A* CE is present in up to 100% of transcripts, suggesting that expression of functional *UNC13A* is greatly reduced, which could impact normal synaptic transmission.

TDP-43 loss induces hundreds of splicing changes, a number of which have also been detected in patient brains. However, it has remained unclear whether these events – even those that occur in crucial neuronal genes – contribute to disease pathogenesis. That genetic variation influencing the *UNC13A* CE inclusion can lead to changes in ALS/FTD susceptibility and progression strongly supports *UNC13A* downregulation to be one of the critical consequences of TDP-43 loss of function. Excitingly, *UNC13A* provides a generalizable therapeutic target for 97% of ALS and approximately half of FTD cases. These findings are also of interest to other neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease and chronic traumatic encephalopathy, in which TDP-43 depletion is also observed in a significant fraction of cases.

Declarations

Acknowledgments:

We thank Frédéric Allain for the His-tagged TDP-43 plasmid, Cristiana Stuani for guidance on TDP-43 purification, and Martina Hallegger for guidance on TDP-43 iCLIP.

Funding:

UK Medical Research Council [MR/M008606/1 and MR/S006508/1] (PF)

UK Motor Neurone Disease Association (PF)

Rosetrees Trust (PF,AG)

UCLH NIHR Biomedical Research Centre (PF)

4 Year Wellcome Trust Studentship (OGW)

European Union's Horizon 2020 research and innovation programme (835300-RNPdynamics) (JU)

Cancer Research UK (FC001002) (JU)

UK Medical Research Council (FC001002) (JU)

Wellcome Trust (FC001002) (JU)

Collaborative Center for X-linked Dystonia-Parkinsonism (WCL, EMCF)

Intramural Research Program of the National Institutes of Neurological Disorders and Stroke, NIH, Bethesda, MD (MEW,SEH)

Chan Zuckerberg Initiative (MEW)

The Robert Packard Center for ALS Research (MEW)

E.B. is funded by AriSLA PathensTDP project (EB)

Wolfson Foundation (AB)

Brightfocus Foundation postdoctoral research fellowship (SEH)

Wellcome Trust Investigator Award (107116/Z/15/Z) (GS)

UK Dementia Research Institute Foundation award (UKDRI-1005) (GS)

Alzheimer's Research UK senior fellowship (TL)

Alzheimers Society (AG)

UKRI Future Leaders Fellowship (MR/T042184/1) (MS)

Author contributions:

Conceptualization: ALB,OGW,MJK,SEH,JH,MEW,PF

Data curation: ALB,OGW,MZ,SBS

Formal analysis: ALB,OGW,MJK,MZ,SBS,AB

Funding acquisition: PF,MEW,EB

Investigation: ALB,OGW,MJK,SEH,MZ,FCYL,LM,YAQ,SBS,AB,WCL,AG

Methodology: ALB,OGW,MJK,SEH,JH,MEW,PF

Project administration: PF,MEW

Resources: HP,TL,EB

Software: ALB,OGW,MZ,SBS,JH

Supervision: PF,MEW,JH,JU,MS,TR,TL,EMCF,GS

Visualization: ALB,OGW,MJK,WCL

Writing – original draft: ALB,OGW,MJK,MEW,PF

Writing – review & editing: SEH,WCL,EB,JU,JH

Competing interests:

ALB, OGW, MJK, SEH, MEW and PF declare competing financial interest. A patent application related to this work has been filed.

Data and materials availability:

Analysis code and data to reproduce figures available:
https://github.com/frattalab/unc13a_cryptic_splicing/

RNA-Seq Data for i3Neurons, SH-SY5Y and SK-N-DZ[a] are available through the European Nucleotide Archive (ENA) under accession PRJEB42763.

Public data was obtained from Gene Expression Omnibus (GEO): iPSC MNs (Klim et al., 2019)-GSE121569, SK-N-DZ[b]-GSE97262, and FACS-sorted frontal cortex neuronal nuclei-GSE126543.

Riboseq: E-MTAB-10235.

Targeted RNA seq: E-MTAB-10237

Minigene iCLIP: E-MTAB-10297

NYGC ALS Consortium RNA-seq: RNA-Seq data generated through the NYGC ALS Consortium in this study can be accessed via the NCBI's GEO database (GEO GSE137810, GSE124439, GSE116622, and GSE153960). All RNA-Seq data generated by 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. To request immediate access to new and ongoing data generated by the NYGC ALS Consortium and for samples provided through the Target ALS Postmortem Core, complete a genetic data request form at ALSData@nygenome.org.

NYGC ALS Consortium Whole Genome Seq: to be released later with companion manuscript.

References

1. R. Ferrari, D. Kapogiannis, E. D. Huey, P. Momeni, FTD and ALS: a tale of two diseases. *Curr. Alzheimer Res.* **8**, 273–294 (2011).
2. P. Couratier, P. Corcia, G. Lautrette, M. Nicol, B. Marin, ALS and frontotemporal dementia belong to a common disease spectrum. *Rev. Neurol. (Paris)*. **173**, 273–279 (2017).
3. A.-L. Ji, X. Zhang, W.-W. Chen, W.-J. Huang, Genetics insight into the amyotrophic lateral sclerosis/frontotemporal dementia spectrum. *J. Med. Genet.* **54**, 145–154 (2017).
4. M. A. van Es, J. H. Veldink, C. G. J. Saris, H. M. Blauw, P. W. J. van Vught, A. Birve, R. Lemmens, H. J. Schelhaas, E. J. N. Groen, M. H. B. Huisman, A. J. van der Kooi, M. de Visser, C. Dahlberg, K. Estrada, F. Rivadeneira, A. Hofman, M. J. Zwarts, P. T. C. van Doormaal, D. Rujescu, E. Strengman, I. Giegling, P. Muglia, B. Tomik, A. Slowik, A. G. Uitterlinden, C. Hendrich, S. Waibel, T. Meyer, A. C. Ludolph, J. D. Glass, S. Purcell, S. Cichon, M. M. Nöthen, H.-E. Wichmann, S. Schreiber, S. H. H. M. Vermeulen, L. A. Kiemeney, J. H. J. Wokke, S. Cronin, R. L. McLaughlin, O. Hardiman, K. Fumoto, R. J. Pasterkamp, V. Meininger, J. Melki, P. N. Leigh, C. E. Shaw, J. E. Landers, A. Al-Chalabi, R. H. Brown, W. Robberecht, P. M. Andersen, R. A. Ophoff, L. H. van den Berg, Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat. Genet.* **41**, 1083–1087 (2009).
5. M. Neumann, D. M. Sampathu, L. K. Kwong, A. C. Truax, M. C. Micsenyi, T. T. Chou, J. Bruce, T. Schuck, M. Grossman, C. M. Clark, L. F. McCluskey, B. L. Miller, E. Masliah, I. R. Mackenzie, H. Feldman, W. Feiden, H. A. Kretzschmar, J. Q. Trojanowski, V. M.-Y. Lee, Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. **314**, 130–133 (2006).
6. J. Humphrey, W. Emmett, P. Fratta, A. M. Isaacs, V. Plagnol, Quantitative analysis of cryptic splicing associated with TDP-43 depletion. *BMC Med. Genomics*. **10**, 38 (2017).
7. J. P. Ling, O. Pletnikova, J. C. Troncoso, P. C. Wong, TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. **349**, 650–655 (2015).
8. J. R. Klim, L. A. Williams, F. Limone, I. Guerra San Juan, B. N. Davis-Dusenbery, D. A. Mordes, A. Burberry, M. J. Steinbaugh, K. K. Gamage, R. Kirchner, R. Moccia, S. H. Cassel, K. Chen, B. J. Wainger,

C. J. Woolf, K. Eggan, ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat. Neurosci.* **22**, 167–179 (2019).

9. Z. Melamed, J. Lopez-Erauskin, M. W. Baughn, O. Zhang, K. Drenner, Y. Sun, F. Freyermuth, M. A. McMahon, M. S. Beccari, J. Artates, T. Ohkubo, M. Rodriguez, N. Lin, D. Wu, C. F. Bennett, F. Rigo, S. Da Cruz, J. Ravits, C. Lagier-Tourenne, D. W. Cleveland, Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat. Neurosci.* **22**, 180–190 (2019).
10. M. Prudencio, J. Humphrey, S. Pickles, A.-L. Brown, S. E. Hill, J. Kachergus, J. Shi, M. Heckman, M. Spiegel, C. Cook, Truncated stathmin-2 is a marker of TDP-43 pathology in frontotemporal dementia. *J. Clin. Invest.* (2020).
11. M. S. Fernandopulle, R. Prestil, C. Grunseich, C. Wang, L. Gan, M. E. Ward, Transcription Factor–Mediated Differentiation of Human iPSCs into Neurons. *Curr. Protoc. Cell Biol.* **79**, e51 (2018).
12. R. Tian, M. A. Gachechiladze, C. H. Ludwig, M. T. Laurie, J. Y. Hong, D. Nathaniel, A. V. Prabhu, M. S. Fernandopulle, R. Patel, M. Abshari, M. E. Ward, M. Kampmann, CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron*. **104**, 239-255.e12 (2019).
13. C. Wang, M. E. Ward, R. Chen, K. Liu, T. E. Tracy, X. Chen, M. Xie, P. D. Sohn, C. Ludwig, A. Meyer-Franke, C. M. Karch, S. Ding, L. Gan, Scalable Production of iPSC-Derived Human Neurons to Identify Tau-Lowering Compounds by High-Content Screening. *Stem Cell Rep.* **9**, 1221–1233 (2017).
14. F. P. Diekstra, P. W. J. van Vught, W. van Rheenen, M. Koppers, R. J. Pasterkamp, M. A. van Es, H. J. Schelhaas, M. de Visser, W. Robberecht, P. Van Damme, P. M. Andersen, L. H. van den Berg, J. H. Veldink, *Neurobiol. Aging*, in press, doi:10.1016/j.neurobiolaging.2011.10.029.
15. F. P. Diekstra, V. M. Van Deerlin, J. C. van Swieten, A. Al-Chalabi, A. C. Ludolph, J. H. Weishaupt, O. Hardiman, J. E. Landers, R. H. Brown, M. A. van Es, R. J. Pasterkamp, M. Koppers, P. M. Andersen, K. Estrada, F. Rivadeneira, A. Hofman, A. G. Uitterlinden, P. van Damme, J. Melki, V. Meininger, A. Shatunov, C. E. Shaw, P. N. Leigh, P. J. Shaw, K. E. Morrison, I. Fogh, A. Chiò, B. J. Traynor, D. Czell, M. Weber, P. Heutink, P. I. W. de Bakker, V. Silani, W. Robberecht, L. H. van den Berg, J. H. Veldink, C9orf72 and UNC13A are shared risk loci for ALS and FTD: a genome-wide meta-analysis. *Ann. Neurol.* **76**, 120–133 (2014).
16. B. Gaastra, A. Shatunov, S. Pulit, A. R. Jones, W. Sproviero, A. Gillett, Z. Chen, J. Kirby, I. Fogh, J. F. Powell, P. N. Leigh, K. E. Morrison, P. J. Shaw, C. E. Shaw, L. H. van den Berg, J. H. Veldink, C. M. Lewis, A. Al-Chalabi, Rare genetic variation in UNC13A may modify survival in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Front. Degener.* **17**, 593–599 (2016).
17. A. Nicolas, K. P. Kenna, A. E. Renton, N. Ticozzi, F. Faghri, R. Chia, J. A. Dominov, B. J. Kenna, M. A. Nalls, P. Keagle, A. M. Rivera, W. van Rheenen, N. A. Murphy, J. J. F. A. van Vugt, J. T. Geiger, R. A. Van der Spek, H. A. Pliner, Shankaracharya, B. N. Smith, G. Marangi, S. D. Topp, Y. Abramzon, A. S. Gkazi, J. D. Eicher, A. Kenna, G. Mora, A. Calvo, L. Mazzini, N. Riva, J. Mandrioli, C. Caponnetto, S. Battistini, P. Volanti, V. La Bella, F. L. Conforti, G. Borghero, S. Messina, I. L. Simone, F. Trojsi, F. Salvi, F. O. Logullo, S. D'Alfonso, L. Corrado, M. Capasso, L. Ferrucci, C. de A. M. Moreno, S. Kamalakaran, D. B.

Goldstein, A. D. Gitler, T. Harris, R. M. Myers, H. Phatnani, R. L. Musunuri, U. S. Evani, A. Abhyankar, M. C. Zody, J. Kaye, S. Finkbeiner, S. K. Wyman, A. LeNail, L. Lima, E. Fraenkel, C. N. Svendsen, L. M. Thompson, J. E. Van Eyk, J. D. Berry, T. M. Miller, S. J. Kolb, M. Cudkowicz, E. Baxi, M. Benatar, J. P. Taylor, E. Rampersaud, G. Wu, J. Wuu, G. Lauria, F. Verde, I. Fogh, C. Tiloca, G. P. Comi, G. Sorarù, C. Cereda, P. Corcia, H. Laaksovirta, L. Myllykangas, L. Jansson, M. Valori, J. Ealing, H. Hamdalla, S. Rollinson, S. Pickering-Brown, R. W. Orrell, K. C. Sidle, A. Malaspina, J. Hardy, A. B. Singleton, J. O. Johnson, S. Arepalli, P. C. Sapp, D. McKenna-Yasek, M. Polak, S. Asress, S. Al-Sarraj, A. King, C. Troakes, C. Vance, J. de Belleroche, F. Baas, A. L. M. A. ten Asbroek, J. L. Muñoz-Blanco, D. G. Hernandez, J. Ding, J. R. Gibbs, S. W. Scholz, M. K. Floeter, R. H. Campbell, F. Landi, R. Bowser, S. M. Pulst, J. M. Ravits, D. J. L. MacGowan, J. Kirby, E. P. Pioro, R. Pamphlett, J. Broach, G. Gerhard, T. L. Dunckley, C. B. Brady, N. W. Kowall, J. C. Troncoso, I. Le Ber, K. Mouzat, S. Lumbroso, T. D. Heiman-Patterson, F. Kamel, L. Van Den Bosch, R. H. Baloh, T. M. Strom, T. Meitinger, A. Shatunov, K. R. Van Eijk, M. de Carvalho, M. Kooyman, B. Middelkoop, M. Moisse, R. L. McLaughlin, M. A. Van Es, M. Weber, K. B. Boylan, M. Van Blitterswijk, R. Rademakers, K. E. Morrison, A. N. Basak, J. S. Mora, V. E. Drory, P. J. Shaw, M. R. Turner, K. Talbot, O. Hardiman, K. L. Williams, J. A. Fifita, G. A. Nicholson, I. P. Blair, G. A. Rouleau, J. Esteban-Pérez, A. García-Redondo, A. Al-Chalabi, E. Rogaeva, L. Zinman, L. W. Ostrow, N. J. Maragakis, J. D. Rothstein, Z. Simmons, J. Cooper-Knock, A. Brice, S. A. Goutman, E. L. Feldman, S. B. Gibson, F. Taroni, A. Ratti, C. Gellera, P. Van Damme, W. Robberecht, P. Fratta, M. Sabatelli, C. Lunetta, A. C. Ludolph, P. M. Andersen, J. H. Weishaupt, W. Camu, J. Q. Trojanowski, V. M. Van Deerlin, R. H. Brown, L. H. van den Berg, J. H. Veldink, M. B. Harms, J. D. Glass, D. J. Stone, P. Tienari, V. Silani, A. Chiò, C. E. Shaw, B. J. Traynor, J. E. Landers, F. O. Logullo, I. Simone, G. Logroscino, F. Salvi, I. Bartolomei, G. Borghero, M. R. Murru, E. Costantino, C. Pani, R. Puddu, C. Caredda, V. Piras, S. Tranquilli, S. Cuccu, D. Corongiu, M. Melis, A. Milia, F. Marrosu, M. G. Marrosu, G. Floris, A. Cannas, S. Tranquilli, M. Capasso, C. Caponnetto, G. Mancardi, P. Origone, P. Mandich, F. L. Conforti, S. Cavallaro, G. Mora, K. Marinou, R. Sideri, S. Penco, L. Mosca, C. Lunetta, G. L. Pinter, M. Corbo, N. Riva, P. Carrera, P. Volanti, J. Mandrioli, N. Fini, A. Fasano, L. Tremolizzo, A. Arosio, C. Ferrarese, F. Trojsi, G. Tedeschi, M. R. Monsurrò, G. Piccirillo, C. Femiano, A. Ticca, E. Ortu, V. La Bella, R. Spataro, T. Colletti, M. Sabatelli, M. Zollino, A. Conte, M. Luigetti, S. Lattante, G. Marangi, M. Santarelli, A. Petrucci, M. Pugliatti, A. Pirisi, L. D. Parish, P. Occhineri, F. Giannini, S. Battistini, C. Ricci, M. Benigni, T. B. Cau, D. Loi, A. Calvo, C. Moglia, M. Brunetti, M. Barberis, G. Restagno, F. Casale, G. Marrali, G. Fuda, I. Ossola, S. Cammarosano, A. Canosa, A. Ilardi, U. Manera, M. Grassano, R. Tanel, F. Pisano, M. B. Harms, D. B. Goldstein, N. A. Shneider, S. Goutman, Z. Simmons, T. M. Miller, S. Chandran, S. Pal, G. Manousakis, S. H. Appel, E. Simpson, L. Wang, R. H. Baloh, S. Gibson, R. Bedlack, D. Lacomis, D. Sareen, A. Sherman, L. Bruijn, M. Penny, A. S. Allen, S. Appel, R. H. Baloh, R. S. Bedlack, B. E. Boone, R. Brown, J. P. Carulli, A. Chesi, W. K. Chung, E. T. Cirulli, G. M. Cooper, J. Couthouis, A. G. Day-Williams, P. A. Dion, S. Gibson, A. D. Gitler, J. D. Glass, D. B. Goldstein, Y. Han, M. B. Harms, T. Harris, S. D. Hayes, A. L. Jones, J. Keebler, B. J. Krueger, B. N. Lasseigne, S. E. Levy, Y.-F. Lu, T. Maniatis, D. McKenna-Yasek, T. M. Miller, R. M. Myers, S. Petrovski, S. M. Pulst, A. R. Raphael, J. M. Ravits, Z. Ren, G. A. Rouleau, P. C. Sapp, N. A. Shneider, E. Simpson, K. B. Sims, J. F. Staropoli, L. L. Waite, Q. Wang, J. R. Wimbish, W. W. Xin, H. Phatnani, J. Kwan, D. Sareen, J. R. Broach, Z. Simmons,

X. Arcila-Londono, E. B. Lee, V. M. Van Deerlin, N. A. Shneider, E. Fraenkel, L. W. Ostrow, F. Baas, N. Zaitlen, J. D. Berry, A. Malaspina, P. Fratta, G. A. Cox, L. M. Thompson, S. Finkbeiner, E. Dardiotis, T. M. Miller, S. Chandran, S. Pal, E. Hornstein, D. J. MacGowan, T. Heiman-Patterson, M. G. Hammell, N. A. Patsopoulos, J. Dubnau, A. Nath, J. Kaye, S. Finkbeiner, S. Wyman, A. LeNail, L. Lima, E. Fraenkel, J. D. Rothstein, C. N. Svendsen, L. M. Thompson, J. Van Eyk, N. J. Maragakis, J. D. Berry, J. D. Glass, T. M. Miller, S. J. Kolb, R. H. Baloh, M. Cudkowicz, E. Baxi, M. Benatar, J. P. Taylor, G. Wu, E. Rampersaud, J. Wuu, R. Rademakers, S. Züchner, R. Schule, J. McCauley, S. Hussain, A. Cooley, M. Wallace, C. Clayman, R. Barohn, J. Statland, J. Ravits, A. Swenson, C. Jackson, J. Trivedi, S. Khan, J. Katz, L. Jenkins, T. Burns, K. Gwathmey, J. Caress, C. McMillan, L. Elman, E. Pioro, J. Heckmann, Y. So, D. Walk, S. Maiser, J. Zhang, V. Silani, N. Ticozzi, C. Gellera, A. Ratti, F. Taroni, G. Lauria, F. Verde, I. Fogh, C. Tiloca, G. P. Comi, G. Sorarù, C. Cereda, S. D'Alfonso, L. Corrado, F. De Marchi, S. Corti, M. Ceroni, L. Mazzini, G. Siciliano, M. Filosto, M. Inghilleri, S. Peverelli, C. Colombrita, B. Poletti, L. Maderna, R. Del Bo, S. Gagliardi, G. Querin, C. Bertolin, V. Pensato, B. Castellotti, W. Camu, K. Mouzat, S. Lumbroso, P. Corcia, V. Meininger, G. Besson, E. Lagrange, P. Clavelou, N. Guy, P. Couratier, P. Vourch, V. Danel, E. Bernard, G. Lemasson, A. Al Kheifat, A. Al-Chalabi, P. Andersen, A. N. Basak, I. P. Blair, A. Chio, J. Cooper-Knock, P. Corcia, P. Couratier, M. de Carvalho, A. Dekker, V. Drory, A. G. Redondo, M. Gotkine, O. Hardiman, W. Hide, A. Iacoangeli, J. Glass, K. Kenna, M. Kiernan, M. Kooyman, J. Landers, R. McLaughlin, B. Middelkoop, J. Mill, M. M. Neto, M. Moisse, J. M. Pardina, K. Morrison, S. Newhouse, S. Pinto, S. Pulit, W. Robberecht, A. Shatunov, P. Shaw, C. Shaw, V. Silani, W. Sproviero, G. Tazelaar, N. Ticozzi, P. van Damme, L. van den Berg, R. van der Spek, K. van Eijk, M. van Es, W. van Rheenen, J. van Vugt, J. Veldink, M. Weber, K. L. Williams, M. Zatz, D. C. Bauer, N. A. Twine, Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron*. **97**, 1268-1283.e6 (2018).

18. K. Placek, G. M. Baer, L. Elman, L. McCluskey, L. Hennessy, P. M. Ferraro, E. B. Lee, V. M.-Y. Lee, J. Q. Trojanowski, V. M. Van Deerlin, M. Grossman, D. J. Irwin, C. T. McMillan, UNC13A polymorphism contributes to frontotemporal disease in sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging*. **73**, 190–199 (2019).

19. C. Pottier, Y. Ren, R. B. Perkerson, M. Baker, G. D. Jenkins, M. van Blitterswijk, M. DeJesus-Hernandez, J. G. J. van Rooij, M. E. Murray, E. Christopher, S. K. McDonnell, Z. Fogarty, A. Batzler, S. Tian, C. T. Vicente, B. Matchett, A. M. Karydas, G.-Y. R. Hsiung, H. Seelaar, M. O. Mol, E. C. Finger, C. Graff, L. Öijerstedt, M. Neumann, P. Heutink, M. Synofzik, C. Wilke, J. Prudlo, P. Rizzu, J. Simon-Sanchez, D. Edbauer, S. Roeber, J. Diehl-Schmid, B. M. Evers, A. King, M. M. Mesulam, S. Weintraub, C. Geula, K. F. Bieniek, L. Petrucelli, G. L. Ahern, E. M. Reiman, B. K. Woodruff, R. J. Caselli, E. D. Huey, M. R. Farlow, J. Grafman, S. Mead, L. T. Grinberg, S. Spina, M. Grossman, D. J. Irwin, E. B. Lee, E. Suh, J. Snowden, D. Mann, N. Ertekin-Taner, R. J. Uitti, Z. K. Wszolek, K. A. Josephs, J. E. Parisi, D. S. Knopman, R. C. Petersen, J. R. Hodges, O. Piguet, E. G. Geier, J. S. Yokoyama, R. A. Rissman, E. Rogaeva, J. Keith, L. Zinman, M. C. Tartaglia, N. J. Cairns, C. Cruchaga, B. Ghetti, J. Kofler, O. L. Lopez, T. G. Beach, T. Arzberger, J. Herms, L. S. Honig, J. P. Vonsattel, G. M. Halliday, J. B. Kwok, C. L. White, M. Gearing, J. Glass, S. Rollinson, S. Pickering-Brown, J. D. Rohrer, J. Q. Trojanowski, V. Van Deerlin, E. H. Bigio, C. Troakes, S. Al-Sarraj, Y. Asmann, B. L. Miller, N. R. Graff-Radford, B. F. Boeve, W. W. Seeley, I. R. A. Mackenzie, J. C. van Swieten, D. W. Dickson, J. M. Biernacka, R. Rademakers, Genome-wide analyses

as part of the international FTLD-TDP whole-genome sequencing consortium reveals novel disease risk factors and increases support for immune dysfunction in FTLD. *Acta Neuropathol. (Berl.)*. **137**, 879–899 (2019).

20. B. Yang, H. Jiang, F. Wang, S. Li, C. Wu, J. Bao, Y. Zhu, Z. Xu, B. Liu, H. Ren, X. Yang, UNC13A variant rs12608932 is associated with increased risk of amyotrophic lateral sclerosis and reduced patient survival: a meta-analysis. *Neurol. Sci.* **40**, 2293–2302 (2019).
21. R. P. A. van Eijk, M. J. C. Eijkemans, S. Nikolakopoulos, M. D. Jansen, H.-J. Westeneng, K. R. van Eijk, R. A. A. van der Spek, J. J. F. A. van Vugt, S. Piepers, G.-J. Groeneveld, J. H. Veldink, L. H. van den Berg, M. A. van Es, Pharmacogenetic interactions in amyotrophic lateral sclerosis: a step closer to a cure? *Pharmacogenomics J.* **20**, 220–226 (2020).
22. J. R. Tollervey, T. Curk, B. Rogelj, M. Briese, M. Cereda, M. Kayikci, T. Hortobágyi, A. L. Nishimura, V. Župunski, R. Patani, S. Chandran, G. Rot, B. Zupan, C. E. Shaw, J. Ule, Characterising the RNA targets and position-dependent splicing regulation by TDP-43; implications for neurodegenerative diseases. *Nat. Neurosci.* **14**, 452–458 (2011).
23. C. Appocher, F. Mohagheghi, S. Cappelli, C. Stuani, M. Romano, F. Feiguin, E. Buratti, Major hnRNP proteins act as general TDP-43 functional modifiers both in Drosophila and human neuronal cells. *Nucleic Acids Res.* **45**, 8026–8045 (2017).
24. E. Y. Liu, J. Russ, C. P. Cali, J. M. Phan, A. Amlie-Wolf, E. B. Lee Correspondence, Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. *CellReports*. **27**, 1409-1421.e6 (2019).
25. J. R. Burrell, G. M. Halliday, J. J. Kril, L. M. Ittner, J. Götz, M. C. Kiernan, J. R. Hodges, The frontotemporal dementia-motor neuron disease continuum. *The Lancet*. **388**, 919–931 (2016).
26. D. Ray, H. Kazan, K. B. Cook, M. T. Weirauch, H. S. Najafabadi, X. Li, S. Gueroussov, M. Albu, H. Zheng, A. Yang, H. Na, M. Irimia, L. H. Matzat, R. K. Dale, S. A. Smith, C. A. Yarosh, S. M. Kelly, B. Nabet, D. Mecenas, W. Li, R. S. Laishram, M. Qiao, H. D. Lipshitz, F. Piano, A. H. Corbett, R. P. Carstens, B. J. Frey, R. A. Anderson, K. W. Lynch, L. O. F. Penalva, E. P. Lei, A. G. Fraser, B. J. Blencowe, Q. D. Morris, T. R. Hughes, A compendium of RNA-binding motifs for decoding gene regulation. *Nature*. **499**, 172–177 (2013).
27. E. H. Corder, A. M. Saunders, W. J. Strittmatter, D. E. Schmechel, P. C. Gaskell, G. W. Small, A. D. Roses, J. L. Haines, M. A. Pericak-Vance, Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. **261**, 921–923 (1993).
28. J. S. Dittman, Unc13: a multifunctional synaptic marvel. *Curr. Opin. Neurobiol.* **57**, 17–25 (2019).
29. F. Varoqueaux, A. Sigler, J.-S. Rhee, N. Brose, C. Enk, K. Reim, C. Rosenmund, Total arrest of spontaneous and evoked synaptic transmission but normal synaptogenesis in the absence of Munc13-mediated vesicle priming. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9037–9042 (2002).
30. F. Varoqueaux, M. S. Sons, J. J. Plomp, N. Brose, Aberrant Morphology and Residual Transmitter Release at the Munc13-Deficient Mouse Neuromuscular Synapse. *Mol. Cell. Biol.* **25**, 5973–5984 (2005).

31. L. A. Gilbert, M. A. Horlbeck, B. Adamson, J. E. Villalta, Y. Chen, E. H. Whitehead, C. Guimaraes, B. Panning, H. L. Ploegh, M. C. Bassik, L. S. Qi, M. Kampmann, J. S. Weissman, Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. **159**, 647–661 (2014).
32. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, Sequence analysis STAR: ultrafast universal RNA-seq aligner. **29**, 15–21 (2013).
33. A. Frankish, M. Diekhans, A.-M. Ferreira, R. Johnson, I. Jungreis, J. Loveland, J. M. Mudge, C. Sisu, J. Wright, J. Armstrong, I. Barnes, A. Berry, A. Bignell, S. Carbonell Sala, J. Chrast, F. Cunningham, T. Di Domenico, S. Donaldson, I. T. Fiddes, C. García Girón, J. M. Gonzalez, T. Grego, M. Hardy, T. Hourlier, T. Hunt, O. G. Izuogu, J. Lagarde, F. J. Martin, L. Martínez, S. Mohanan, P. Muir, F. C. P. Navarro, A. Parker, B. Pei, F. Pozo, M. Ruffier, B. M. Schmitt, E. Stapleton, M.-M. Suner, I. Sycheva, B. Uszczynska-Ratajczak, J. Xu, A. Yates, D. Zerbino, Y. Zhang, B. Aken, J. S. Choudhary, M. Gerstein, R. Guigó, T. J. P. Hubbard, M. Kellis, B. Paten, A. Reymond, M. L. Tress, P. Flicek, GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
34. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. **30**, 923–930 (2014).
35. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
36. F. Mölder, K. P. Jablonski, B. Letcher, M. B. Hall, C. H. Tomkins-Tinch, V. Sochat, J. Forster, S. Lee, S. O. Twardziok, A. Kanitz, A. Wilm, M. Holtgrewe, S. Rahmann, S. Nahnsen, J. Köster, Sustainable data analysis with Snakemake. *F1000Research*. **10**, 33 (2021).
37. J. Vaquero-Garcia, A. Barrera, M. R. Gazzara, J. González-Vallinas, N. F. Lahens, J. B. Hogenesch, K. W. Lynch, Y. Barash, A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife*. **5**, e11752 (2016).
38. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**, 841–842 (2010).
39. R. Middleton, D. Gao, A. Thomas, B. Singh, A. Au, J. J.-L. Wong, A. Bomane, B. Cosson, E. Eyras, J. E. J. Rasko, W. Ritchie, IRFinder: assessing the impact of intron retention on mammalian gene expression. *Genome Biol.* **18**, 51 (2017).
40. K. J. Livak, T. D. Schmittgen, Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*. **25**, 402–408 (2001).
41. A. P. Pereverzev, N. G. Gurskaya, G. V. Ermakova, E. I. Kudryavtseva, N. M. Markina, A. A. Kotlobay, S. A. Lukyanov, A. G. Zaraisky, K. A. Lukyanov, Method for quantitative analysis of nonsense-mediated mRNA decay at the single cell level. *Sci. Rep.* **5**, 7729 (2015).
42. J. Humphrey, N. Birsa, C. Milioto, M. McLaughlin, A. M. Ule, D. Robaldo, A. B. Eberle, R. Kräuchi, M. Bentham, A.-L. Brown, S. Jarvis, C. Bodo, M. G. Garone, A. Devoy, G. Soraru, A. Rosa, I. Bozzoni, E. M. C. Fisher, O. Mühlemann, G. Schiavo, M.-D. Ruepp, A. M. Isaacs, V. Plagnol, P. Fratta, FUS ALS-causative mutations impair FUS autoregulation and splicing factor networks through intron retention. *Nucleic Acids Res.* **48**, 6889–6905 (2020).

43. L. Blazquez, W. Emmett, R. Faraway, J. M. B. Pineda, S. Bajew, A. Gohr, N. Haberman, C. R. Sibley, R. K. Bradley, M. Irimia, J. Ule, Exon Junction Complex Shapes the Transcriptome by Repressing Recursive Splicing. *Mol. Cell*. **72**, 496-509.e9 (2018).
44. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. **9**, 357–359 (2012).
45. T. Smith, A. Heger, I. Sudbery, UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
46. A. Buniello, J. A. L. MacArthur, M. Cerezo, L. W. Harris, J. Hayhurst, C. Malangone, A. McMahon, J. Morales, E. Mountjoy, E. Sollis, D. Suveges, O. Vrousgou, P. L. Whetzel, R. Amode, J. A. Guillen, H. S. Riat, S. J. Trevanion, P. Hall, H. Junkins, P. Flicek, T. Burdett, L. A. Hindorff, F. Cunningham, H. Parkinson, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
47. R. J. Pruim, R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Gliedt, M. Boehnke, G. R. Abecasis, C. J. Willer, LocusZoom: regional visualization of genome-wide association scan results. *Bioinforma. Oxf. Engl.* **26**, 2336–2337 (2010).
48. 1000 Genomes Project Consortium, A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, J. O. Korbel, J. L. Marchini, S. McCarthy, G. A. McVean, G. R. Abecasis, A global reference for human genetic variation. *Nature*. **526**, 68–74 (2015).
49. O. H. Tam, N. V. Rozhkov, R. Shaw, J. Ravits, J. Dubnau, M. Gale, H. Correspondence, Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *Cell Rep.* **29** (2019), doi:10.1016/j.celrep.2019.09.066.
50. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25**, 2078–2079 (2009).
51. Picard toolkit. *Broad Inst. GitHub Repos.* (2019) (available at http://broadinstitute.github.io/picard/).
52. K. C. Cotto, Y.-Y. Feng, A. Ramu, Z. L. Skidmore, J. Kunisaki, M. Richters, S. Freshour, Y. Lin, W. C. Chapman, R. Uppaluri, R. Govindan, O. L. Griffith, M. Griffith, RegTools: Integrated analysis of genomic and transcriptomic data for the discovery of splicing variants in cancer. *bioRxiv*, 436634 (2021).
53. Y. I. Li, D. A. Knowles, J. Humphrey, A. N. Barbeira, S. P. Dickinson, H. K. Im, J. K. Pritchard, Annotation-free quantification of RNA splicing using LeafCutter. *Nat. Genet.* **50**, 151–158 (2018).
54. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio* (2013) (available at http://arxiv.org/abs/1303.3997).
55. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

56. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

57. P. J. Lukavsky, D. Daujotyte, J. R. Tollervey, J. Ule, C. Stuani, E. Buratti, F. E. Baralle, F. F. Damberger, F. H.-T. Allain, Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449 (2013).

Figures

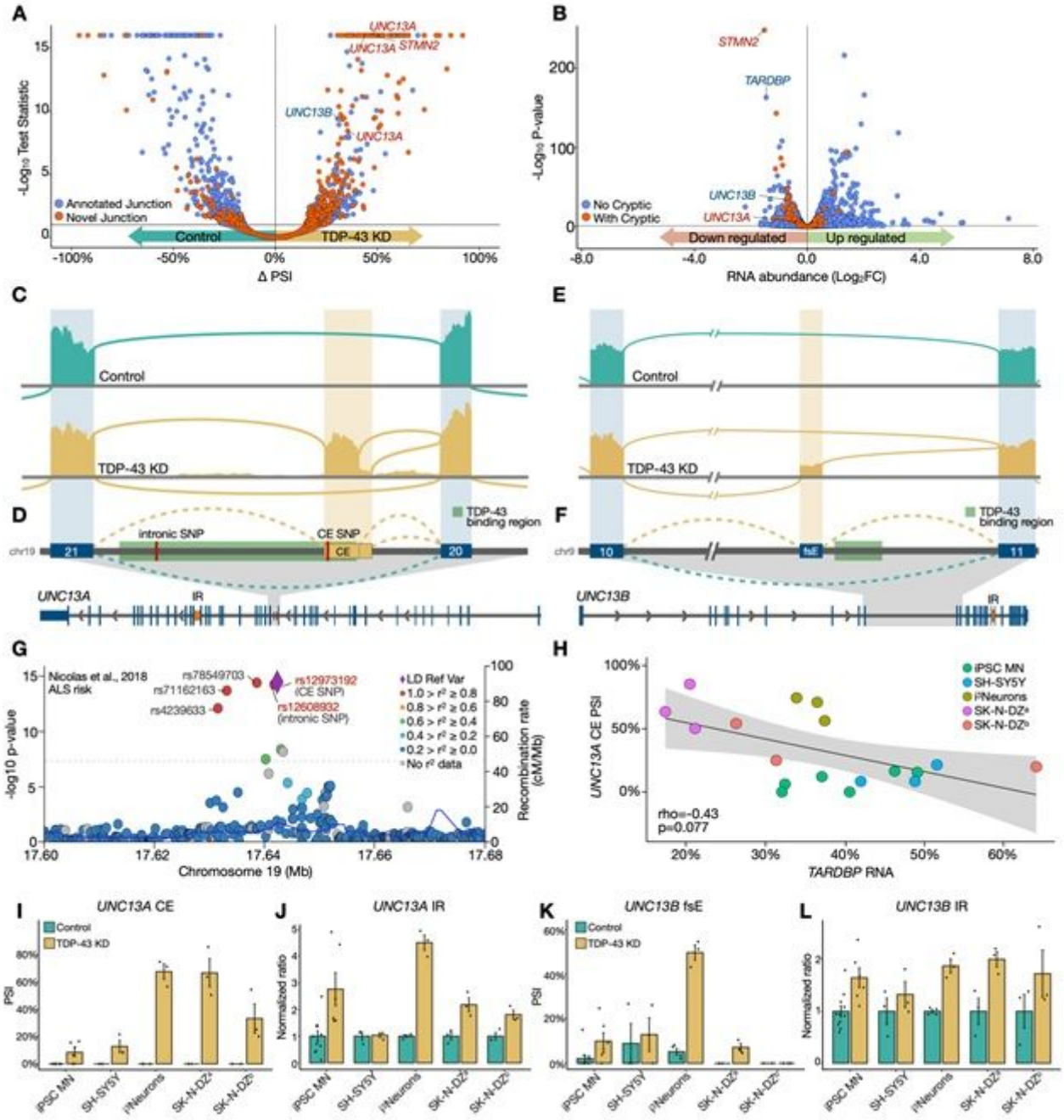

Figure 1

TDP-43 depletion in neurons leads to altered splicing in synaptic genes UNC13A and UNC13B. (A) Differential splicing and (B) expression in control (N=4) and CRISPRi TDP-43 depleted (N=3) iPSC-derived cortical-like i3Neurons. Each point denotes a splice junction (A) or gene (B). (C) Sashimi plots showing cryptic exon (CE) inclusion between exons 20 and 21 of UNC13A upon TDP-43 knockdown (KD). (D,F) Schematics showing intron retention (IR, lower schematic, orange), TDP-43 binding region (22)(green), and two ALS/FTLD associated SNPs (red). (E) Sashimi plot of UNC13B showing inclusion of the frameshifting exon (fsE) upon TDP-43 KD. (G) LocusZoom plot of the UNC13A locus in the latest ALS GWAS. Lead SNP rs12973192 plotted as purple diamond, other SNPs coloured by linkage disequilibrium with rs12973192 in European individuals from 1000 Genomes. (H) Correlation between relative TARDBP RNA and UNC13A CE PSI across five TDP-43 knockdown datasets (I,K) PSI of TDP-43 regulated splicing in UNC13A and UNC13B across neuronal datasets. (J,L) Intron retention ratio of TDP-43 regulated retained introns in UNC13A and UNC13B across neuronal datasets.

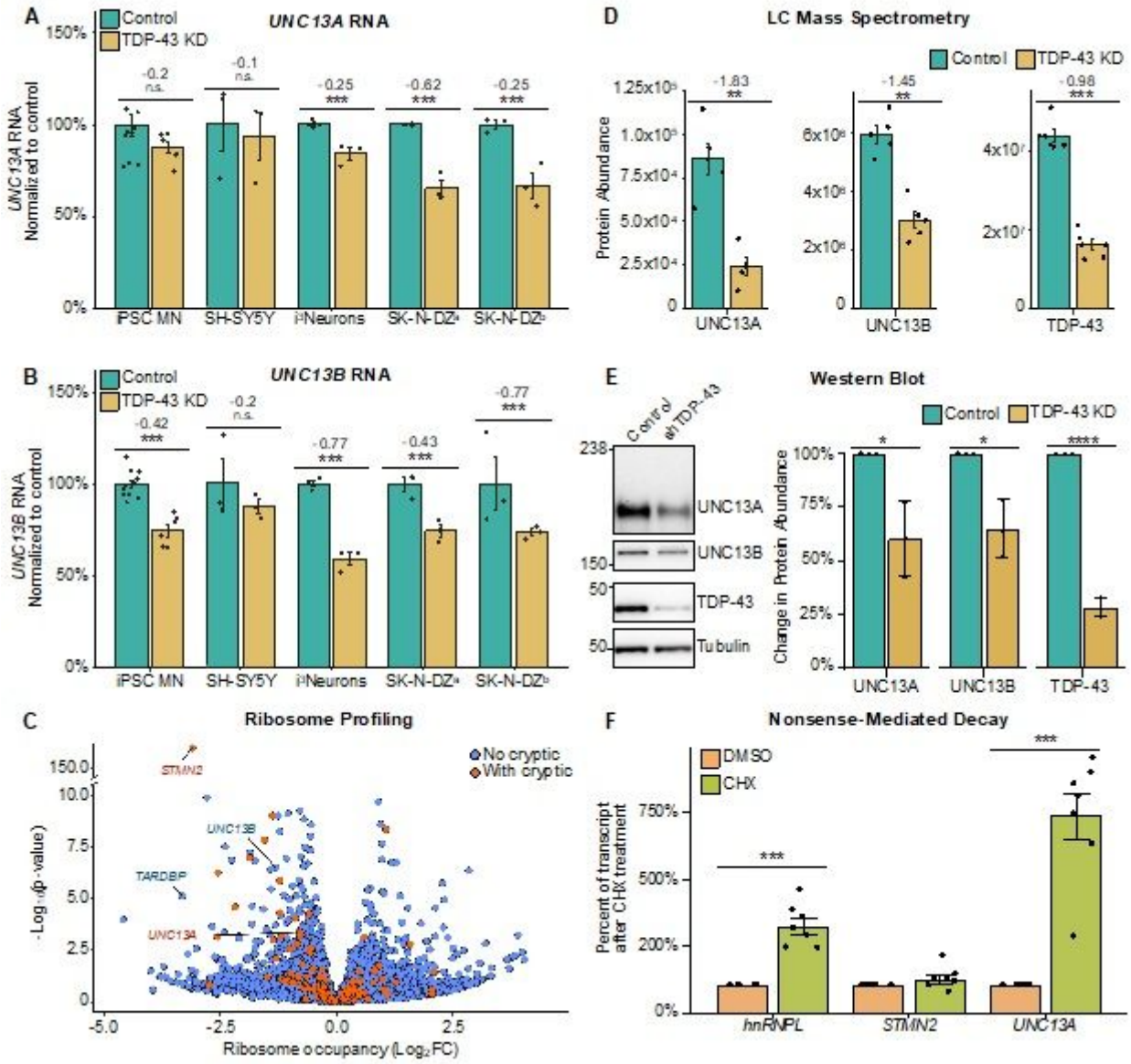

Figure 2

UNC13A and UNC13B are downregulated after TDP-43 knockdown due to the production of NMD-sensitive transcripts. Relative gene expression levels for UNC13A (A) and UNC13B (B) after TDP-43 knockdown across neuronal cell lines. Normalized RNA counts are shown as relative to control mean. Numbers show log2 fold change calculated by DESeq2. Significance shown as adjusted p-values from DESeq2. (C) Ribosome profiling of TDP-43 knockdown in i3Neurons shows reduction in ribosome occupancy of STMN2, UNC13A and UNC13B transcripts. (D) Mass spectrometry-based proteomic analysis shows reduction in protein abundance of UNC13A, UNC13B and TDP-43 upon TDP-43 knockdown in i3Neurons. Numbers refer to log2 fold change of unique peptide fragments, P-values from Wilcoxon test. (E) Western blot analysis of protein lysates from untreated and TDP-43 knockdown SH-SY5Y cells show a significant reduction in UNC13A and UNC13B proteins levels after TDP-43 depletion. Graphs represent the means ± S.E., N=3, One sample t-test, (F) Transcript expression upon CHX treatment suggests UNC13A but not STMN2, are sensitive to nonsense-mediated decay. HNRNPL (heterogeneous nuclear ribonucleoprotein L) is a positive control. Significance levels reported as * (p<0.05) ** (p<0.01) *** (p<0.001) **** (p <0.0001).

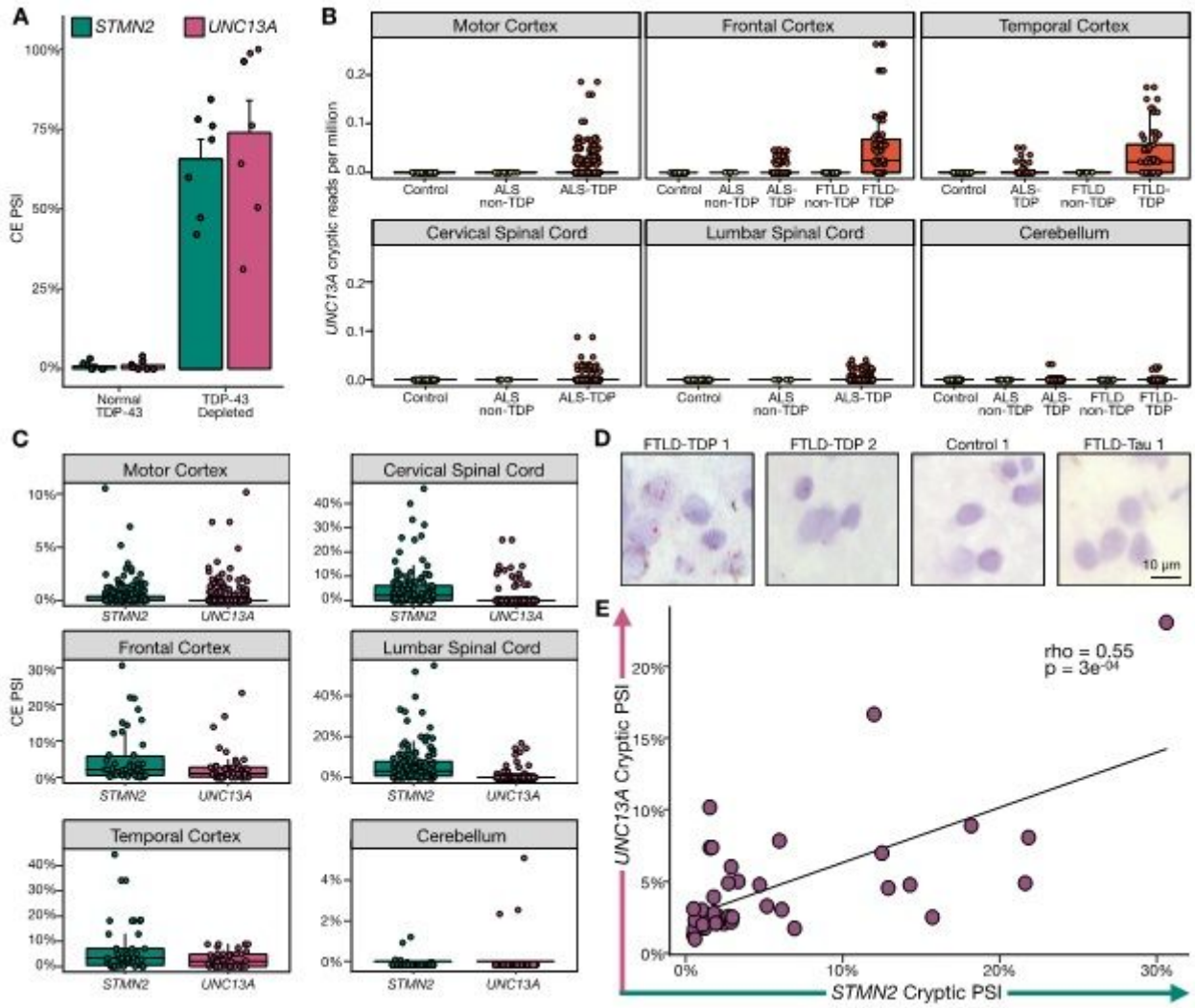

Figure 3

UNC13A CE is highly expressed in ALS/FTLD patient tissue and correlates with known markers of TDP-43 loss of function. (A) UNC13A and STMN2 CE expression in ALS/FTLD patient frontal cortex neuronal nuclei from24 sorted according to the expression of nuclear TDP-43. (B) UNC13A CE expression in bulk RNA-seq from NYGC ALS Consortium normalized by library size across disease and tissue samples. ALS cases stratified by mutation status, FTLD cases stratified by pathological subtype. (C) CE expression throughout ALS/FTLD-TDP cases across tissue (D) BaseScope detection of UNC13A CE (red foci) in FTLD-TDP but not control or FTLD-Tau frontal cortex samples. (E) Correlation in ALS/FTLD-TDP cortex between UNC13A and STMN2 CE PSI in patients with at least 30 spliced reads across the CE locus.

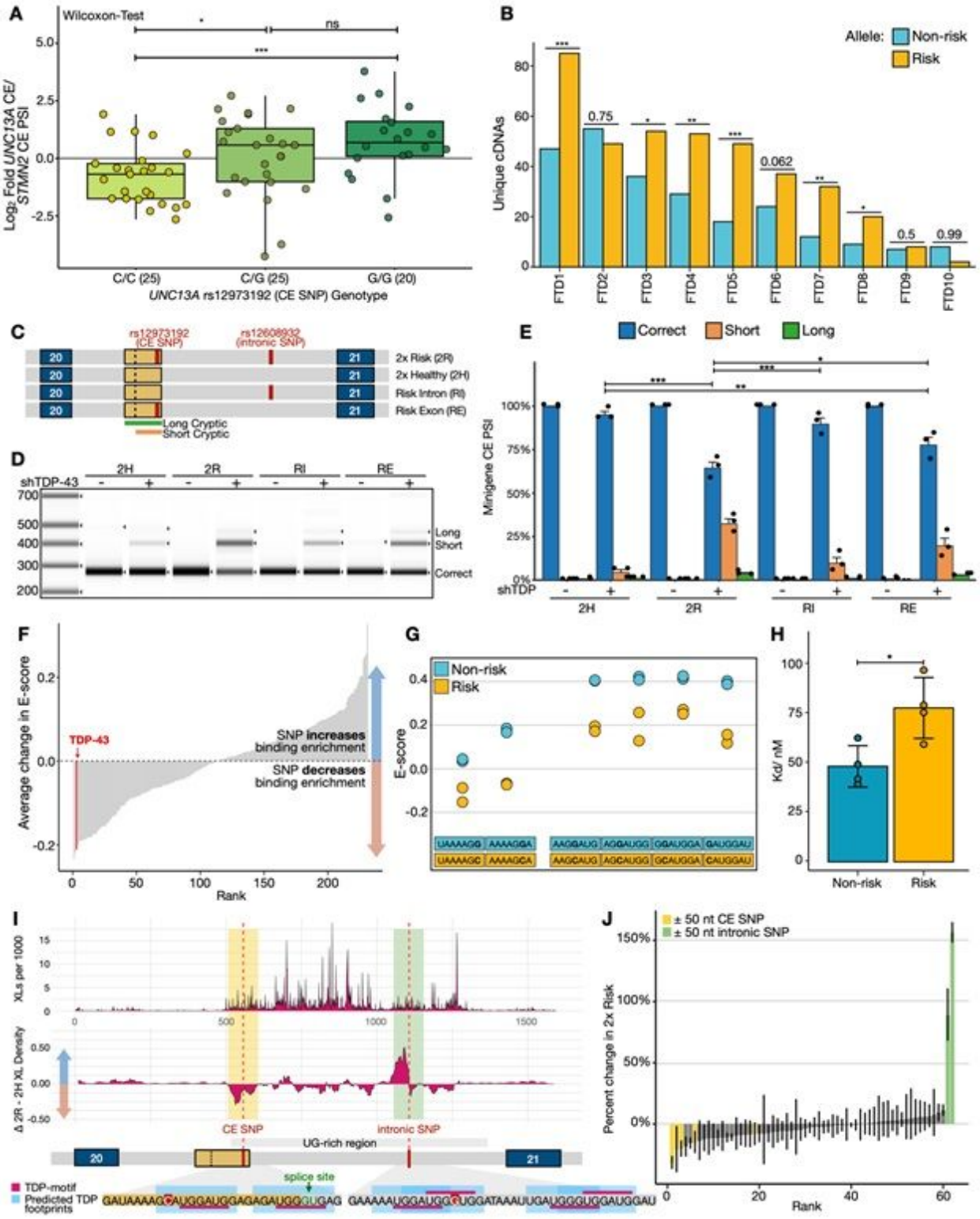

Figure 4

UNC13A ALS/FTD risk variants enhance UNC13A CE splicing in patients and in vitro by altering TDP-43 pre-mRNA binding. (A) Ratio UNC13A / STMN2 CE PSI, split by genotype for UNC13A risk alleles. (B) Unique cDNAs from targeted RNA-seq in ten CE SNP heterozygous FTLD-TDP patients. p-values from single-tailed binomial tests. FTD1, 5, and 7 are C9orf72 hexanucleotide repeat carriers (C) Illustration of UNC13A minigenes containing exon 20, intron 20, and exon 21 with both risk SNPs (2R), both healthy

SNPs (2H), or risk SNP in CE (RE) or intron (RI). (D) Representative image of RT-PCR products from UNC13A minigenes in SH-SY5Y ± TDP-43 KD. (E) Quantification of (D) plotted as means ± S.E. N=3, One-way ANOVA analysis; (F) Average change in E-value (measure of binding enrichment) across proteins for heptamers containing risk/healthy CE SNP allele; red - TDP-43. (G) Each CE SNP heptamer's TDP-43 E-value. (H) Binding affinities between TDP-43 and 14-nt RNA containing the healthy or risk sequence measured by ITC; 4 replicates. (I) TDP-43 iCLIP of 2R and 2H minigenes: top - average crosslink density; bottom - average density change 2R - 2H (rolling window = 20 nt, units = crosslinks per 1,000). Cartoon - predicted TDP-43 binding footprints (UGNNUG motif). (J) Fractional changes at iCLIP peaks for 2R versus 2H minigene (mean and 75% confidence interval shown). Peaks that are within 50nt of each SNP are highlighted. *** ($p<0.001$) ** ($p<0.01$) * ($p<0.05$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.pdf
- DataS1.csv
- DataS2.csv
- DataS3.csv
- DataS4.xlsx