## Title page

# Counter-regulation of RNA stability by UPF1 and TDP43


## Author list

Nicolas Gomez[1], Caroline Hsieh[1], Xingli Li[2], Megan Dykstra[3], Jacob Waksmacki[2], Christopher Altheim[2], Yoel Bechar[2], Joseph Klim[4], Benjamin Zaepful[5], Jeffrey Rothstein[6], Elizabeth EM Tank[2], Sami J Barmada[1,2]

## Affiliations

[1]Graduate Program in Cell and Molecular Biology, University of Michigan, Ann Arbor, MI

[2]Department of Neurology, University of Michigan, Ann Arbor, MI

[3]Neuroscience Graduate Program, University of Michigan, Ann Arbor, MI

[4]Aquinnah Pharmaceuticals, Cambridge, MA

[5]Department of Neurology, University of California, San Diego, CA

[6]Department of Neurology, Johns Hopkins University, Baltimore, MD

*To whom correspondence should be addressed

Sami Barmada

University of Michigan

Department of Neurology

109 Zina Pitcher Place, 4019 BSRB

Ann Arbor, MI 48109

## Abstract

RNA quality control is crucial for proper regulation of gene expression. Disruption of nonsense mediated mRNA decay (NMD), the primary RNA decay pathway responsible for the degradation of transcripts containing premature termination codons (PTCs), can disrupt development and lead to multiple diseases in humans and other animals. Similarly, therapies targeting NMD may have applications in hematological, neoplastic and neurological disorders. As such, tools capable of accurately quantifying NMD status could be invaluable for investigations of disease pathogenesis and biomarker identification. Toward this end, we assemble, validate, and apply a next-generation sequencing approach (NMDq) for identifying and measuring the abundance of PTC-containing transcripts. After validating NMDq performance and confirming its utility for tracking RNA surveillance, we apply it to determine pathway activity in two neurodegenerative diseases, amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) characterized by RNA misprocessing and abnormal RNA stability. Despite the genetic and pathologic evidence implicating dysfunctional RNA metabolism, and NMD in particular, in these conditions, we detected no significant differences in PTC-encoding transcripts in ALS models or disease. Contrary to expectations, overexpression of the master NMD regulator UPF1 had little effect on the clearance of transcripts with PTCs, but rather restored RNA homeostasis through differential use and decay of alternatively poly-adenylated isoforms. Together, these data suggest that canonical NMD is not a significant contributor to ALS/FTD pathogenesis, and that UPF1 promotes neuronal survival by regulating transcripts with abnormally long 3’UTRs.

## Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disease characterized by the progressive loss of upper and lower motor neurons[1,2]. How paralysis manifests and the extent to which cognitive function is involved varies significantly between individuals[3]. Despite this clinical and genetic heterogeneity, 95% of individuals with ALS exhibit cytoplasmic accumulation of TDP43 (transactive response element DNA/RNA binding protein, 43 kDa) in affected neurons and glia[4,5]. TDP43 is a ubiquitously expressed splicing factor that is primarily localized to the nucleus in healthy cells[6,7]. Mutations in the gene encoding TDP43 (*TARDBP*) and several similar RNA-binding proteins (RBPs) (e.g. FUS, HNRNPA2B1, MATR3) cause familial ALS as well as the related, often comorbid disease frontotemporal dementia (FTD)[8–10].

In cellular and animal model systems, TDP43 overexpression results not only in neurodegeneration but also in TDP43 cytosolic mislocalization and aggregation, recapitulating key pathologic changes seen in humans[11–18]. We and others determined that overexpression of the RNA helicase up-frameshift 1 (UPF1), and to a lesser extent its binding partner UPF2, mitigates toxicity in both *in vivo* and *in vitro* models of ALS/FTD associated with TDP43 and FUS accumulation[12,13,19]. More recently, a similar neuroprotective effect was observed in *Drosophila* and neuron models of familial ALS/FTD due to *C9ORF72* hexanucleotide repeat expansions[20–22], raising the possibility that UPF1 could act broadly to prevent neurodegeneration in ALS and FTD.

UPF1 is an obligate component of nonsense-mediated decay (NMD), a vital RNA surveillance pathway responsible for degrading premature termination codon (PTC)-containing transcripts that would otherwise produce truncated and potentially deleterious proteins[23,24]. These data are consistent with the detection of several new NMD substrates associated with TDP43 dysfunction[25–30], as well as abnormal RNA stability in cells derived from ALS/FTD patients[31]. Little is known, however, about NMD activity in ALS/FTD models or patients, and whether any detected deficiencies are selective for these disorders.

Despite the importance of NMD to both human health and disease, quantifying NMD in biological systems remains an area of contention. Previous strategies employing overexpression-based reporter systems are limited in both generalizability and tractability[32,33], particularly *in vivo*. Endogenous biomarkers of NMD activity identified in next-generation sequencing (NGS)-based screens, on the other hand, have taken an inherently limited “gene-centric” view of NMD that fails to take into account the vulnerability of differentially spliced transcript isoforms for a single gene[21,34]. Consequently, deriving NMD biomarkers from differential gene expression may suffer from low sensitivity and specificity.

To overcome these and other limitations we present an analytical pipeline, NMDq, that leverages available tools (STAR, StringTie2, and IsoformSwitchAnalyzeR), to identify novel, assayable, and highly sensitive NMD substrates harboring PTCs. We then use both a global and candidate-based approach to characterize the status of NMD in ALS/FTD models and tissues. Finally, we explore alternative RNA

surveillance pathways through which UPF1 achieves neuroprotection, providing the basis for novel disease-modifying therapeutics for patients with ALS and FTD.

## Results

### NMDq RNA-seq analytic pipeline leverages available bioinformatic methods to quantify global NMD burden

Prior methods for assaying the clearance of PTC-containing NMD substrates have been limited in either isoform specificity or model generalizability. To overcome these limitations and quantify NMD at the transcript level in a variety of model systems, we implemented a modified version of a widely used RNA-seq analytic pipeline (**Fig. 1a**). Paired-end reads were aligned to the hg39 genome assembly using the splice-aware aligner STAR[35]. A *de novo* transcriptome was assembled using StringTie2[36,37] and the GENCODE v38 annotation as a reference. The dataset-specific *de novo* transcriptome was then used to quantify aligned reads. Both annotated and novel transcripts were assessed for PTCs using the R package IsoformSwitchAnalyzeR[38], which implements DEXseq[39] to perform differential expression analysis of quantified transcripts.

As a proof of concept for the NMD quantification (NMDq) pipeline, we first applied it to a previous dataset wherein HeLa cells were treated with siRNA targeting essential NMD-pathway members SMG6 and SMG7, or non-targeting siRNA[34]. In each dataset, an approximately equal number of novel NTC- and PTC-containing isoforms were discovered through *de novo* transcriptome assembly, reducing the likelihood of bias towards either NMD class (**Fig. 1b**).

To determine the impact of unannotated gene products on measurements of NMD, we focused specifically on representative genes such as *DNAJC17* that display both NTC- and PTC-containing isoforms (**Fig. 1c-e**). Without the addition of novel isoforms, SMG6/7 knockdown appeared to selectively downregulate NMD-insensitive (i.e. NTC-containing) isoforms. In contrast, the inclusion of novel isoforms using NMDq demonstrated a pronounced effect of SMG6/7 knockdown on the abundance of NMD-sensitive (i.e. PTC-containing) isoforms, consistent with impaired surveillance of NMD substrates. To evaluate the cumulative consequence of NMD deficiency, isoform abundances were aggregated by NMD-class (**Fig. 1f**). Estimating the absolute NMD burden in this manner, we detected a >2-fold increase in PTC-containing isoforms upon SMG6/7 knockdown.

Alternative splicing and RNA-surveillance are stochastic processes capable of producing myriad low-abundance isoforms[40,41]. As a result, direct comparison of scarce NMD substrates (TPM<5) between groups can be particularly vulnerable to noise. To overcome this, isoforms were aggregated by NMD-class on a per-gene basis and used to determine differential NMD-class usage (DNU), representing the fraction of isoforms subject to NMD for each gene. DNU allows for direct comparison of the contribution of NMD to individual gene abundance, providing insight into those with similar sequence-dependent splicing potential in

different biological contexts. In HeLa cells, DNU analysis identified 2929 genes with significantly increased NMD-class usage upon SMG6/7 knockdown, and 183 with increased NMD-class usage in the control group (log2 fold-change ≥ 1.5, adjusted p-value ≤ 0.05) (**Fig. 1g**). Together, these data confirm the utility of NMDq for measuring both global and gene-specific changes in the abundance of PTC-containing NMD substrates.

**NMDq is sensitive to both genetic- and pharmacologic-driven impairments in NMD in a variety of cell-based model systems**

Due to the cell-specific nature of alternative splicing[42] and NMD activity[43], characterizing NMD in a universal, platform-independent manner can be challenging. Furthermore, because NMD involves the coordination of several distinct components, functional NMD inhibition can be achieved through several different mechanisms. To assess the sensitivity of NMDq across cell types and modes of NMD inhibition, we applied the pipeline to six distinct RNA-seq datasets (**Table 1**). NMDq was able to detect not just large changes, but also subtle differences in PTC-containing isoform abundances in HeLa cells, HEK293T cells, and human induced pluripotent stem cell (iPSC)-derived motor neurons (iMNs) when NMD was impaired by genetic or pharmacologic means (**Fig. 2**). Whereas differences in annotated isoforms alone suggested that the primary consequence of NMD inhibition was a loss of NTC-containing isoforms (**Fig. 2a-c, g-i**), the inclusion of novel unannotated isoforms demonstrated an unambiguous upregulation of PTC-containing transcripts in NMD-deficient conditions, consistent with direct inhibition of this pathway. In these experiments, UPF1 knockdown produced only a modest increase in NMD-substrates (**Fig. 2g**); however, measurement of NMD efficiency in this experiment could be complicated by difficulties in achieving adequate UPF1 knockdown, and the intrinsic toxicity of UPF1 loss of function[44].

Although DNU analysis can be helpful for gauging the contribution of NMD to the homeostasis of individual genes, we also sought to develop a singular measurement that encompasses NMD activity as a whole, enabling rapid comparison of NMD function across experimental conditions. We therefore summed the PTC-containing fraction of total reads and normalized this value to control, generating a global estimate of NMD burden (GNB). Total PTC+ abundance was greater in all NMD-deficient groups compared to respective controls (**Fig. 2g-h**, filled bars), consistent with effective inhibition of NMD in each condition. Because the majority of genes encode NMD-insensitive products (**Fig. 1b**), we performed a secondary, focused analysis of NMD substrates by limiting the PTC-containing fraction to intrinsically NMD-sensitive genes (INS), or those involved in statistically significant (q≤0.05) differential NMD-class switching (**Fig. 2g-j**, empty bars). The cumulative abundance of PTC+ isoforms encoded by INS genes was between 2 and 4-fold greater in NMD-deficient groups compared to respective controls. Moreover, many INS genes were common to at least two datasets: 325 genes were found in all immortalized cell datasets, and 59 of these were also detected within iMN datasets, suggesting the possibility of convergent and sensitive markers of NMD impairment. These candidates included several members of the serine-arginine rich splicing factor family, including *SRSF2*, *SRSF6*, *TRA2B* (also known as *SRSF10*). Each of these proteins regulates its expression by enhancing the

inclusion of a PTC-containing “poison exon” within their cognate mRNA transcript, leading to its degradation by NMD[45].

**NMDq identifies assayable targets of canonical NMD**

To define universal markers of NMD impairment, we selected a series of novel PTC-containing NMD targets identified by NMDq that were expressed broadly, abundantly, and consistently across cell types. These conserved transcripts were compared by abundance within immortalized cells and iMNs (**Fig. 3a, Supp. Fig. 1a**). Nine highly abundant isoforms demonstrating prominent upregulation upon NMD inhibition were selected for RT-PCR-based validation (red points in **Fig. 3b, e**). HEK293T cells and iMNs were treated with either DMSO or the SMG1 inhibitor 11j[46] for 48h and harvested for protein and RNA. The relative RNA abundance of selected PTC-containing isoforms were then compared between treatments. In keeping with the ability of NMDq to highlight conserved NMD substrates, all candidates showed significant upregulation upon treatment with 11j (**Fig. 3c, f**). Notably, novel NMD substrates identified by NMDq were considerably more sensitive as indicators of NMD impairment than previously published targets (*ATF3*, *ATF4*, and *ERN1*)[47]. *BAG1* (BAG family molecular chaperone regulator 1), a gene with a previously-identified NMD-sensitive transcript isoform[48], showed a >12-fold increase upon 11j treatment in both HEK293T cells and iMNs. *SRSF2* exhibits highly-conserved poison exon sequences that are essential for the regulation of its expression via a process that requires NMD[45]. We observed a marked increase in poison exon usage in 11j-treated iMNs compared to DMSO-treated (representative examples given in **Supp. Fig. 1c**) suggesting that poison exon inclusion as a mode of NMD-mediated gene regulation is more widespread than previously appreciated.To determine if changes in transcript abundance upon NMD inhibition affect protein levels, we assayed one of the most abundant and broadly-expressed candidate NMD substrates, *SAT1*. In doing so, we observed a ~1.5-fold increase in protein expression in 11j-treated HEK293T cells compared to controls (**Supp. Fig. 1b**), demonstrating that changes in NMD efficiency can have significant consequences at the protein level.

NMDq and all RNA-seq-based assays measure the steady-state abundance of transcripts and are therefore limited in their ability to accurately estimate the activity of NMD or other pathways that selectively degrade RNA. To determine whether the observed changes in the abundance of candidate NMD substrates were due to RNA turnover, we turned to roadblock-qPCR, a technique enabling the direct measurement of RNA clearance[49]. As expected, *GAPDH* and exogenous firefly luciferase RNA were stable and unaffected by NMD inhibition (**Fig. 3d**), while polyadenylated 18S ribosomal RNA was highly unstable[50], and degraded rapidly through an NMD-independent process. Importantly, NMD-substrates identified by NMDq (*BAG1*, *SAT1*, and *SRSF2*) were unstable in DMSO-treated cells, and each transcript was significantly stabilized upon NMD inhibition, confirming their degradation through NMD.

**Global NMD burden is not consistently elevated in ALS models and human tissue**

Although RNA metabolism is significantly altered in ALS/FTD[31] patient samples, it remains unclear whether abnormalities in canonical NMD are responsible for these differences. Furthermore, overexpression of UPF1—an essential NMD factor—rescues phenotypes of disease in several ALS/FTD model systems[12,13,19–22]. These observations raise the possibility that NMD could be fundamentally impaired in ALS/FTD, and that UPF1 overexpression relieves this impairment. To address this question, we used NMDq to assess PTC-containing transcripts in several iMN models of ALS/FTD, including disease due to the *C9ORF72* hexanucleotide repeat expansion[51,52], *TARDBP* mutation (M337V)[8,53], TDP43 overexpression or knockdown[12], and sporadic onset ALS (sALS) (**Table 2**). *C9ORF72* iMNs showed a subtle and nonsignificant decrease in GNB compared to control iMNs (**Fig. 4a**). These observations are consistent with previous studies suggesting that pathogenic *C9ORF72* mutations may in fact stimulate NMD activity[21,22]. iMNs derived from sALS iPSCs also showed a modest trend towards lower GNB compared to control iMNs. Conversely, TDP43 overexpression and knockdown, in addition to the M337V *TARDBP* mutation, all elicited mild and nonsignificant increases in GNB compared to controls (**Fig. 4b**). The impact of TDP43 overexpression on NMD activity was neither dose-dependent nor mutation-related: high-level TDP43 expression via lentiviral transduction had the same effect as a single copy insertion of TDP43, and overexpression of TDP43(M337V) was indistinguishable from TDP43(WT). Together, these data argue against NMD downregulation in patient-derived neuronal models of ALS/FTD.

Analogous trends were observed in human tissues (**Fig. 4c**). Using data from three separate studies of *C9ORF72*-linked ALS, we noted no significant change in the GNB to suggest changes in overall NMD activity. In sALS, we detected only a modest reduction in GNB, similar to what we originally found in sALS iMNs (**Fig. 4b**). A separate analysis of RNA-seq data from Huntington's disease brain[54] also showed no substantial differences from controls. Due to variability in sample preparation and sequencing methods across ALS datasets, we sought to validate these findings by specifically focusing on the abundance of NMDq-derived NMD-substrates within each of the iMN and human tissue datasets described in **Fig. 4a-c**. In each case, we observed little to no change in NMD substrates (**Fig. 4d**). As a positive control, we detected a robust increase in the abundance of the same set of NMD substrates in iMNs after the application of the NMD inhibitor 11j. These results suggest that NMD is not significantly affected in ALS/FTD or iMN models of these diseases.

The above studies are limited by data availability and the scarcity of samples from individuals with familial ALS/FD. Therefore, to examine the possibility that rare pathogenic mutations associated with ALS impact NMD, we expressed a broad array of ALS/FTD-associated proteins bearing disease-associated mutations in HEK293T cells, and measured the abundance of three NMD candidate substrates identified by NMDq (*BAG1*, *SAT1*, and *SRSF2*) by RT-qPCR (**Fig. 4e**). As expected, application of 11j resulted in a marked increase in all three candidates. However, among a panel of TDP43, FUS, SOD1, C9ORF72, and MATR3 variants, only FUS overexpression resulted in significant upregulation of NMD substrates. These results are consistent with previous studies suggesting that pathogenic FUS variants impair NMD. Even so, the majority of pathogenic variants displayed little to no effect on the same NMD substrates, arguing against a significant impact of TDP43, SOD1, C9ORF72, and MATR3 variants on NMD activity.

**UPF1 and TDP43 counter-regulate the essential arginine biosynthetic enzyme ASS1**

In the absence of detectable NMD impairment in ALS disease models, we broadened our search for molecular pathways responsible for UPF1-mediated neuroprotection. We reanalyzed our lentivirus-treated iMN RNA-seq data, focusing on genes showing significant changes upon TDP43 overexpression that were subsequently rescued by co-expression of UPF1. Several candidates displaying counter regulation by TDP43 and UPF1 emerged from this analysis, including transcripts encoding ly6/PLAUR domain-containing protein 1 (LYPD1), laminin subunit alpha-4 (LAMA4), and argininosuccinate synthetase 1(ASS1) (**Fig. 5a**). LYPD1 is a neurotransmitter receptor binding protein of unknown function that is enriched in Von Economo neurons[55] and layer 5b cortical neurons that are selectively vulnerable in FTD[56,57] and mouse models of SOD1 ALS[58], respectively. LAMA4 is an extracellular matrix protein of particular importance to neuromuscular junction physiology[59], while ASS1 is an essential enzyme that catalyzes the rate-limiting step in arginine biosynthesis. Mutations affecting ASS1 function or expression cause familial citrulinemia, a pediatric neurological disease characterized by seizure and delayed cognitive development[60]. Notably, rescue by UPF1 was unaffected by 11j for each of these candidates, suggesting a mechanism independent of UPF1 phosphorylation.

Due to its established role in neurologic disease, we pursued *ASS1* as a candidate indicator of UPF1's effects. First, we verified the reciprocal effect of TDP43 and UPF1 on *ASS1* mRNA by RT-qPCR in human iMNs (**Fig. 5b**). To determine if TDP43 directly regulated *ASS1*, we transiently transfected HEK293T cells with either WT or an RNA-binding-deficient (RBD-null) TDP43 variant carrying the F147L and F149L mutations[61,62]. As in iMNs (**Fig. 4a, b**), *ASS1* was downregulated by TDP43 overexpression; this effect required RNA binding, as *ASS1* expression was unperturbed in TDP43 RBD-null expressing cells (**Fig. 5c**). These data suggest that TDP43 directly binds *ASS1* mRNA, consistent with the existence of multiple UG-rich repeats within the ASS1 transcript[63]. To confirm this possibility, we performed a streptavidin pull-down assay of *ASS1* RNA-containing ribonucleoprotein (RNP) particles[64]. UV-crosslinked HEK293T cells were lysed and incubated with biotin-conjugated antisense oligonucleotides (btn-ASO) which were then immobilized and washed on streptavidin-bound iron nanoparticles. We designed btn-ASOs to hybridize with either *ASS1* and confirmed the specificity of these probes via PCR of the target, *ASS1* mRNA (**Fig. 5d**), compared to a scrambled ASO control. We also assessed the composition of ASS1-containing RNPs by SDS-PAGE and Western blotting. *ASS1* btn-ASO eluates contained both TDP43 and UPF1 protein but lacked TUB1A, confirming assay specificity (**Fig. 5e**), and suggesting that *ASS1* mRNA interacts directly with both TDP43 and UPF1.

Cryptic splicing in the absence of functional TDP43 results in downregulation of essential factors such as stathmin-2 (STMN2) in ALS/FTD tissue[26,27]. To determine if *ASS1* is similarly dysregulated by TDP43 *in vivo*, and to validate *ASS1* levels as a potential biomarker of TDP43 activity, we utilized single-molecule fluorescence *in situ* hybridization (smFISH) to examine *ASS1* transcripts in post-mortem spinal cord from individuals with ALS (**Fig. 6a-c**). *STMN2* and *ASS1* transcripts were abundant in spinal neurons from the anterior horn of controls, but both transcripts were effectively depleted from ALS spinal neurons. Thus, similar

to *STMN2*, *ASS1* transcripts are downregulated in ALS spinal neurons. Re-analysis of RNA-seq data from *C9ORF72* ALS/FTD post-mortem tissue[65] confirmed this observation, demonstrating a >50% drop in *ASS1* levels selectively in neurons with TDP43 pathology (**Fig. 6d**). As *ASS1* loss is associated with neurological phenotypes in citrulinemia[60], we asked whether *ASS1* loss-of-function is sufficient for neurotoxicity. For this, we utilized automated microscopy and survival analysis, a technique we optimized for monitoring the survival of rodent primary neurons in several different disease-associated contexts[11,12,61,66,67]. Mixed cortical neurons were isolated from non-transgenic rodents, then transfected with EGFP and an RFP-tagged plasmid encoding either shRNA targeting the coding sequence of *ASS1* or a scrambled sequence. Transfected neurons were followed daily via fluorescence microscopy, and individual neurons were identified by image segmentation software (**Fig. 6e**). Time of death for each cell was determined by morphological criteria (neurite retraction, soma rounding, dissolution or loss of fluorescence) that proved to be highly sensitive indicators of cell death[68]. Taking advantage of the variability of transient transfection in neurons[12,68,69], we stratified neurons by RFP fluorescence (encoded on the same plasmid as *ASS1* shRNA), and selected cells expressing high or low levels of shRNA. Compared to scrambled shRNA, shASS1 high-expressing neurons displayed a significant reduction in ASS1 protein, whereas shASS1 low-expressing neurons exhibited a non-significant trend towards lower ASS1 levels (**Fig. 6f**). We then compared the survival of neurons with low- or high-expression of ASS1 shRNA to those transfected with non-targeting shRNA (**Fig. 6g**). For both low and high-expressing neurons, *ASS1* depletion resulted in a significant increase in the cumulative risk of death compared to controls, suggesting that even slight reductions in *ASS1* expression can be neurotoxic.

Conversely, we asked whether restoring *ASS1* levels could mitigate TDP43-mediated toxicity. We transfected primary neurons with plasmids encoding either EGFP or TDP43-EGFP, together with mApple or ASS1 tagged with mApple, and imaged daily for 10d as before (**Fig. 6h**). As in prior investigations[11,12,61,66,69], TDP43-EGFP overexpression resulted in a significant increase in the cumulative risk of death compared to EGFP alone. Although ASS1 expression had no apparent effect on the survival of EGFP-expressing control neurons, ASS1 over-expression resulted in a modest, yet significant increase in risk of death for TDP43-EGFP-expressing neurons, suggesting that ASS1 restoration is insufficient to rescue TDP43 related toxicity, and in fact can compound the toxicity associated with TDP43 accumulation.

**UPF1 and TDP43 counter-regulate global 3'UTR length and RNA stability**

Independent of its splicing activity, TDP43 can modulate RNA stability by regulating polyA usage and 3'UTR length[27,70,71]. We therefore wondered if changes in the expression of *ASS1* and other genes identified in TDP43 over-expressing iMNs correlate with 3'UTR usage. To do this, we applied DaPars[72] to detect differential use of novel and annotated polyA sites in our lentivirus-treated iMN RNA-seq dataset. Using the mean percentage of Distal polyA site Usage (DPAU), we found 4280 3'UTR-lengthening changes and 267 3'UTR-shortening events in TDP43-overexpressing iMNs, compared to iMNs transduced with EGFP alone (**Fig. 7a**). This is consistent with previous illustrations of proximal alternative polyA sites and shortened 3'UTRs

upon TDP43 knockdown[26,27,71]. Comparing UPF1 and TDP43 co-expressing iMNs to EGFP expressing iMNs, we found only 536 3'UTR-lengthening and 787 3'UTR-shortening events (**Fig. 7b**), confirming the link between 3'UTR length and a UPF1-dependent, PTC-independent mRNA decay mechanism[34,73,74]. To explore whether similar changes were present *in vivo*, we applied DaPars to RNA-seq data from *C9ORF72*-ALS/FTD frontal cortex[75]. These analyses uncovered 335 3'UTR-lengthening and 57 3'UTR-shortening PAU events in *C9ORF72*-ALS/FTD cortex compared to that from healthy controls (**Fig. 7c**). Together these data highlight 3'UTR lengthening is a conserved feature of ALS/FTD pathology, and furthermore suggest that UPF1 expression can partially reverse the observed increases in distal polyA site usage.

As TDP43 and UPF1 induced counter-directional changes in polyA site usage and 3'UTR length, and both factors are intricately connected with RNA stability, we asked whether RNA turnover is likewise affected oppositely by TDP43 and UPF1. To examine RNA clearance in iMNs, we took advantage of metabolic pulse-chase labeling of nascent RNA using thiol(SH)-containing uridine analogs, coupled with RNA-sequencing (SLAMseq)[76]. Lentivirus-treated iMNs were incubated with 4-thiouridine (4sU) for 45m to label actively-transcribed RNA, and then washed with unlabeled uridine (**Fig. 7d**). For each condition, samples were collected immediately after the labeling period as well as after 4h of uridine chase. Purified 4sU-containing RNA was chemically alkylated, leading to mis-interpretation of labeled uridine residues as cytosine by next-generation sequencing. Transitions are then quantified in each read to determine the approximate fraction of "new" and "old" RNA as well as half-life on a gene-by-gene basis using GRAND-SLAM[77].

To determine the selectivity of labeling, we first examined the relative rate of transitions by nucleobase (**Fig. 7e**). In the majority of 4sU-labeled samples, we observed the expected T-to-C transitions at nearly 20-fold the rate of unlabeled iMNs, consistent with efficient labeling. The rate of off-target labeling, on the other hand, was less than 3-fold in the majority of cases. We next examined the effect of UPF1 and TDP43 overexpression on global RNA stability (**Fig. 7f**). UPF1 overexpression induced a leftward shift in the distribution of RNA half-lives compared to EGFP, consistent with broad RNA destabilization and reflective of the canonical function of UPF1 in myriad RNA decay pathways. TDP43 overexpression, on the other hand, induced a rightward shift in RNA half-lives, suggesting overall transcript stabilization. This conflicts with previous data showing a net destabilization of RNA in iPSCs transiently expressing TDP43[31], potentially a result of cell type-specific consequences of TDP43-driven 3'UTR lengthening. That is, whereas 3'UTR shortening is generally stabilizing in most mammalian cell types[78,79], 3'UTR lengthening may uniquely stabilize RNA in neurons[80]. Notably, coexpression of UPF1 and TDP43 yielded a distribution of RNA half-lives that was identical to that of EGFP overexpression, suggesting that UPF1 and TDP43 counter-regulate the stability of the neuronal transcriptome as a whole.

To further understand the determinants of UPF1- and TDP43-driven changes in RNA stability, we first categorized transcripts as direct TDP43 RNA targets or non-targets, based on available TDP43 CLIP-seq data[81,82] (**Fig. 7g**). In doing so, we observed that TDP43 targets were preferentially stabilized by TDP43

overexpression in comparison to non-targets, suggesting a selective effect of TDP43 on RNA half-life driven by direct associations of TDP43 and its target transcripts. Conversely, the destabilizing effect of UPF1 was unselective, consistent with previous findings that the nature of UPF1 target discrimination is largely independent of sequence motifs[73,83,84]. We also asked whether genes demonstrating differential polyA site usage via steady-state RNA-seq were differentially stabilized by TDP43 expression as determined by SLAM-seq. Indeed, genes that exhibited TDP43-driven 3'UTR-lengthening were significantly more stabilized than those that exhibited no change in 3'UTR length or those that displayed 3'UTR-shortening (**Fig. 7h**). Additionally, genes that exhibited TDP43-driven 3'UTR-lengthening were significantly more destabilized by UPF1 coexpression compared to genes showing no change in 3'UTR length or 3'UTR-shortening (**Fig. 7i**). These data suggest that TDP43-induced 3'UTR-lengthening events primarily stabilize gene products, an effect that is reversed by UPF1 expression. For example, the gene encoding vesicle membrane protein synaptogamin (*SYT1*) exhibited 3'UTR-lengthening changes upon TDP43 overexpression and in *C9ORF72*-linked ALS/FTD frontal cortex (**Fig. 7j**), and increased stability in TDP43-overexpressing iMNs. UPF1 co-expression normalized *SYT1* stability in iMNs, however, in conjunction with restored 3'UTR length and polyA site usage (**Fig. 7k**). These data suggest that TDP43 preferentially drives distal polyA site usage, conferring increased stability, and that UPF1 coexpression eliminates 3'UTR-lengthened transcripts through a PTC-independent and 3'UTR length-dependent decay mechanism.

## Discussion

In this study, we explored whether NMD was impaired in ALS and UPF1-treated models of ALS/FTD. To accomplish this we devised a next generation sequencing-approach to quantify the abundance of PTC-containing transcripts in a platform-independent manner. After validating our approach with genetic and pharmacologic models of NMD impairment, we applied our pipeline to ten ALS/FTD RNA-seq datasets as well as one Huntington disease RNA-seq dataset. We found no evidence of elevated PTC-containing transcripts in disease contexts. In fact, several iMN models indicate subtle decreases in NMD-sensitive reads, suggesting subtle NMD enhancement. In the absence of obvious NMD impairment, we expanded our search for PTC-independent mechanisms of UPF1-related neuroprotection. We identified *ASS1* as a TDP43-downregulated and UPF1-normalized gene whose loss of function has critical implications for neuronal health, and whose expression is dramatically reduced in spinal neurons from ALS patients. Finally, we demonstrated that TDP43 and UPF1 have antagonistic effects on steady state 3'UTR length as well as RNA stability in a manner consistent with PTC-independent, long-3'UTR-triggered UPF1 activity. These findings challenge the prevailing PTC-based model of UPF1-mediated neuroprotection and provide the basis for an underappreciated role of UPF1 in selectively clearing transcripts with abnormally long 3'UTRs. Our observations are mirrored by concurrent studies from Alessandrini et al., confirming the accumulation of transcripts with long 3'UTRs in patient-derived samples and human neuron ALS models. Together, these investigations argue for a novel mechanism of neuroprotection by UPF1 involving the clearance of alternatively

polyadenylated transcripts with long 3'UTRs induced by TDP43 dysfunction, a highly prevalent and signature neuropathological event in the majority of those with ALS and FTD.

To identify novel and broadly expressed NMD substrates in an unbiased manner, we implemented a bioinformatics pipeline and applied both global and candidate-based strategies to characterize NMD function in ALS/FTD models and disease. The pipeline, termed NMDq, was constructed from widely available software that, to our knowledge, had not previously been used as presented to explore the pathophysiology of ALS, FTD, or other neurological conditions. To validate the accuracy and biological relevance of NMDq, we first analyzed 6 RNA-seq datasets of NMD-impaired cell systems: 3 from publicly available repositories, 2 from collaborators, and 1 from this study. We also included data from a variety of cell types with both genetically and pharmacologically-driven NMD impairment, ensuring the universality of our approach. We devised two complimentary statistics that effectively summarize NMD activity from sequencing data: global NMD burden (GNB), or the cumulative TPM of transcripts containing PTCs, and differential NMD-class usage (DNU), or the relative preference for PTC-containing transcripts on a gene-by-gene basis. Using GNB, we can draw conclusions about the global capacity of represented cell systems to surveil and metabolize derangements in the transcriptome pertaining specifically to canonical NMD affecting PTC-containing transcripts. In contrast, using DNU we can assay the sensitivities of specific genes to the inhibition of NMD by way of the balance of NMD-sensitive to NMD-insensitive products.

Taking this further, we used a set analysis of DNU-positive genes (genes that showed significant differential NMD-class usage) across all included datasets to identify consistent and reproducible PTC-containing NMD substrates. We used the isoform expression characteristics of NMD-sensitive gene products to identify eight candidates for which we designed isoform-specific and total gene qRT-PCR primers. Several of these candidate substrates, such as *SRSF2*, *SRSF3*, *SRSF6*, and *TRA2B,* belong to a family of serine-arginine splicing factors that autoregulate themselves by directing the inclusion of a PTC-containing "poison" exon[45] (as in **Supp. Fig. 1c**). Other targets, including *BAG1* and *SAT1*, are subject to NMD due to alternative splicing and the inclusion of conserved exons carrying stop codons[48,85]. Several substrates, including *TRIR* and *DDX1*, have no prior association with NMD and are therefore novel targets; the latter of these (*DDX1*), is a neuronal-selective marker of NMD that may be invaluable for future studies of cell type-specific NMD regulation in neurodegenerative disease models.

To validate our candidate NMD-substrates, we implemented both steady-state and nascent qRT-PCR. At steady-state, all candidates showed significantly improved sensitivity to pharmacologic NMD inhibition compared to previously published targets, including *ATF3*, *ATF4*, and *ERN1*[47]. To explore whether these changes were due to a post-transcriptional destabilization event or an increase in transcription, we employed roadblock-qPCR[49], a significantly less toxic and more direct approach to studying RNA stability compared to the most common approach involving transcriptional inhibitors such as actinomycin-D[86]. Our results show that NMD inhibition results in upregulation of *BAG1*, *SAT1*, and *SRSF2* primarily via stabilization of otherwise highly unstable transcripts. Importantly, enhancements in transcript stability were also associated with modest

increases at the protein level for SAT1. Additional studies are warranted to determine the stability of the resulting truncated protein product and the extent to which antibody or epitope selection influences the detection of such shortened or out-of-frame products.

Encouraged by the ability of NMDq to accurately identify changes in NMD activity at the global level as well as for individual transcripts and genes, we applied the pipeline to ten ALS/FTD RNA-seq datasets and one from Huntington's disease. We did not find convincing evidence of NMD dysfunction in either cell-based models or human post-mortem tissue, challenging the prevailing suspicion that UPF1 overexpression prevents neurodegeneration by restoring or enhancing canonical NMD activity. Although growing evidence points towards altered RNA metabolism in ALS models and disease[31,87,88], our data suggest that such impairments are not a consequence of insufficient capacity to degrade or otherwise resolve PTC-containing transcripts. In fact, in datasets relevant to *C9ORF72*-related ALS/FTD as well as sporadic ALS, we noted subtly ***reduced*** NMD-substrate burden consistent with ***increased*** canonical NMD. A notable exception to this global trend of intact NMD function may be the preponderance of NMD-substrates in HEK293 cells overexpressing FUS or mutant variants of FUS. Consistent with previous studies[88], we found both WT and disease-associated FUS mutations elicit small increases in NMD-sensitive gene products of *BAG1*, *SAT1*, and *SRSF2*. We also failed to find any evidence of differentially active, phosphorylated UPF1 in post-mortem tissue from ALS patients. Taken together, these observations argue against global NMD impairment in ALS.

Without clear evidence of NMD dysfunction, specifically with regard to the clearance of PTC-containing transcripts, we expanded our search for mediators of UPF1-associated neuroprotection through conventional, gene-based analyses of RNA-sequencing data from iMNs overexpressing TDP43 and/or UPF1. In the process, we identified several genes that responded robustly and antagonistically to TDP43 and UPF1 overexpression. One such gene, *ASS1*, encodes an essential enzyme that catalyzes the rate-limiting step in arginine biosynthesis. Loss of ASS1 function results in type 1 citrullinemia, a primary metabolic disorder characterized by significant neurologic impairment[60]. Our studies show that TDP43, but not RNA binding-deficient TDP43 (F147L, F149L), drives the downregulation of this vital gene, suggesting that this effect is mediated via direct interaction of TDP43 with the *ASS1* transcript. Biotinylated ASO pull-down assays confirmed this possibility, revealing direct interactions between *ASS1* RNA and both TDP43 as well as UPF1. In keeping with the TDP43-dependent downregulation of *ASS1* we observed in iMNs, we also detected a dramatic reduction in *ASS1* expression within spinal motor neurons of ALS patients, compared to controls. *ASS1* was also significantly downregulated in neurons with TDP43 pathology, as detected by prior studies involving fluorescence-based sorting of neuronal nuclei from *C9ORF72* ALS/FTD post-mortem cortex[65]. Furthermore, longitudinal microscopy and automated survival analysis of primary neurons confirmed the importance of *ASS1* in preventing neuron loss, in accord with previous studies showing that neurons are highly sensitive to both decreased and increased ASS1[60,89,90]. ASS1 overexpression was insufficient to rescue TDP43 toxicity, however. There are at least two explanations for this finding. First, TDP43 regulates thousands of transcripts[81,82], many of which including *ASS1*, are crucial. Therapeutic restoration of *ASS1* in and of itself may not be sufficient to prevent TDP43-driven toxicity. Second, *ASS1* restoration may require finer control of gene

expression than is possible with transient transfection. In murine models of citrullinemia, for example, *ASS1* gene therapy using adeno-associated virus prevents premature death and lowers plasma citrulline, but paradoxically results in arginine depletion and renal failure through unknown mechanisms[90].

To further elucidate the means by which TDP43 and UPF1 counter-regulate gene expression, we quantified differential polyA site usage using DaPars[72]. Prior studies suggested that TDP43 binds to and influences the use of distal polyA sites on nascent RNA, thereby affecting the expression of TDP43 itself[71] as well as many other genes[70]. In accord with these findings, our data demonstrate preferential distal polyA site usage in TDP43-overexpressing iMNs, together with an overall increase in global RNA stability. UPF1, on the other hand, has no established influence on polyA site usage. Instead, it may regulate mature mRNA levels by sensing 3'UTR length and potentiating EJC- and PTC-independent transcript decay[73,84,91–93]. Co-expression of UPF1 prevented TDP43-induced distal polyA site usage, and effectively countered changes in RNA stabilization associated with TDP43 expression. Although additional studies are required to directly interrogate the mechanisms behind UPF1-mediated length-sensing of TDP43-lengthened transcripts, these investigations are in keeping with the consistent ability of UPF1 expression to prevent TDP43-related toxicity in multiple model systems, from yeast to primary and human neurons, flies, and rodents[12,13,19–22].

We also asked whether the changes in polyA site usage associated with TDP43 and UPF1 expression might correlate with changes in global RNA stability. To answer this question, and to identify candidate mediators of UPF1-driven rescue that were not evident from steady-state RNA-seq studies, we employed SLAM-seq[76,77], wherein total RNA is sequenced and estimations of nascent RNA synthesis and stability are derived from C-T transitions induced by 4-thiol-uridine labeling. We found that global RNA stability was reduced by the expression of UPF1 and increased by the expression of TDP43. The observed effect of UPF1 expression on RNA stability is consistent with all canonical and non-canonical functions of UPF1 as an RNA decay factor. TDP43's effects on RNA stability, however, conflict with previous findings in undifferentiated iPSCs, obtained through the related technique BruChase-seq[31]. This discrepancy could be due to a number of factors. First, determinants of RNA stability are cell type-specific[43,94]. A relevant example of this is the differential impact of 3'UTR length on RNA stability in neurons compared to non-neuronal cells. In neurons, 3'UTR lengthening stabilizes transcripts, whereas the opposite relationship between 3'UTR length and RNA stability is observed in non-neuronal cell types[80,95,96]. Second, differences in the techniques used to study RNA stability may account for some of the observed experimental discordance. In BruChase-seq, for example, bromouridine-containing RNA is purified from total RNA via immunoprecipitation using anti-bromouridine antibodies prior to sequencing[97]. SLAM-seq avoids the prohibitively high input requirements for BruChase-seq. The consequences of these technical distinctions have not been directly interrogated, but may stem from differences in cDNA library depth and coverage of modified RNA. In SLAM-seq, transition-containing reads represent only a small fraction of total reads which may result in sample error. In contrast, all reads are bromouridine-containing in BruChase-seq, but unmodified reads are unavailable to provide additional insight into differences in steady-state RNA abundance that precede labeling. Together, these emerging techniques

offer complimentary perspectives on our developing understanding of RNA stability in experimental cell systems.

## Materials and methods

### NMDq

Paired-end RNA-seq reads from publicly accessible datasets were quality and adapter trimmed (TrimGalore v0.6.6, default parameters) and aligned to the human genome (STAR v2.7.3a; GRCh38, Gencode v35) with the optional parameters: "--twopassMode Basic --outSAMtype BAM SortedByCoordinate --outSAMattributes NH HI NM MD AS nM jM jI XS --quantMode GeneCounts"[35]. For each dataset, a de novo transcriptome inclusive of all identified junctions was assembled ("--merge") and BAM files were requantified (StringTie v2.1.4, default parameters)[36,37]. TPMs were calculated and open reading frames (ORFs) were estimated to predict premature termination codons (PTCs) (IsoformSwitchAnalyzeR v1.10.0)[38]. Transcripts were categorized into "NMD" or "non-NMD" classes according to their Ensembl annotated or predicted sensitivity to NMD. TPMs were aggregated class- and gene-wise and summarized with two approaches: 1) Global NMD burden: Group-wise sums of NMD-TPM was as a proportion of total TPM was and normalized to within-dataset controls. 2) Differential NMD-class usage (DNU): Gene-wise ratios of NMD to total transcript TPM were calculated (NMD ratio). NMD ratio was log-transformed and used to fit generalized linear models to gene- and group-wise comparisons (limma v3.44.3)[98]. An empirical Bayes test was used on the resulting model to determine statistical significance. Genes showing ≥1.5 fold change from compared conditions were counted as exhibiting DNU.

### HEK293T cell culture

Human embryonic kidney (HEK) 293T cells were cultured in DMEM (GIBCO), 10% FBS, 100 units/mL Penicillin/Streptomycin at 37°C in 5% CO2. Cells were passaged when they reached 75-80% confluency using Accutase (ThermoFisher #MT25058CI), and transfected using Lipofectamine 2000 (Fisher Scientific #11668019) according to the manufacturer's protocol.

### iMotor neuron differentiation and transduction

iMotor neurons were differentiated as described in Chua et al. (2022)[99]. Briefly, on day 0 of differentiation, neural precursor cells were washed in PBS and incubated in prewarmed accutase (Sigma, #A6964) at 37°C for 8m. Four volumes of E8 media (ThermoFisher, #A1517001) were added to the plate, and the cells were collected and pelleted at 200xg for 5m. The media was aspirated, and the pellet was resuspended in 1mL of

fresh E8 media. Cells were counted using a hemocytometer, diluted, plated at a density of 20,000 cells/mL in E8 media with ROCK inhibitor and incubated at 37°C overnight. On day 1, the cells were placed in N2 media (1x N2 Supplement (Gibco, #17502- 048), 1x NEAA Supplement (Gibco, #11140-050), 10 ng/mL BDNF (Peprotech, #450-02), 10 ng/mL NT3 (Peprotech, #450-03), 0.2 µg/mL laminin (Sigma, #L2020), 2 mg/mL doxycycline (Sigma, #D3447) in E8 media). On day 2, N2 media was replaced with transition media ((1x N2 Supplement, 1x NEAA Supplement, 10 ng/mL BDNF, 10 ng/ml NT3, 0.2 µg/mL laminin, 2 mg/mL doxycycline in half E8 media, half DMEM F12 (Gibco, #11320-033)). On day 3, transition media was replaced with B27 media (1x B27 Supplement (Gibco, #17504-044), 1x Glutamax Supplement (Gibco, #35050-061), 10 ng/mL BDNF, 10 ng/mL NT3, 0.2 µg/mL laminin, and 1x Culture One (Gibco, #A33202-01) in Neurobasal-A (Gibco, #12349-015)). On day 9, cells were transduced with the appropriate virus (prepared by University of Michigan Vector Core) and cells were sustained in the same culture medium for the remainder of the experiment. On day 11, iMNs were treated with 1uM 11j SMG1-inhibitor or an equivalent volume of DMSO. On day 14, iMNs were collected for RNA isolation.

**Next-generation sequencing**

cDNA libraries were prepared from Trizol extracted, DNA-digested samples using the Illumina Stranded mRNA prep kit (Illumina, #20040532). Paired end sequencing was carried out on an Illumina NovaSeq (S4) 300 cycle sequencer at the University of Michigan Advanced Genomics Core. Samples were sequenced at an average depth of 75M reads/sample.

**Roadblock qPCR**

Live HEK293T or iMN cells were metabolically labeled for 6h with the nucleoside analog 4-thiouridine (4sU, final concentration 400uM) in 2h intervals. Following RNA extraction and DNase treatment, incorporated 4sU was modified using the thiol-reactive compound N-ethylmaleimide (NEM, 50ug/ul) in a compatible buffer (125mM Tris-HCl pH 8.0, 2.5mM EDTA, ultrapure water). 0.2ng/ul firefly luciferase mRNA (FL, synthesized in house) was spiked into thiol modified samples. Reverse transcription was performed using the ProtoScript® II First Strand cDNA Synthesis Kit (NEB, #E6560S) and 18-mer oligo dT primers. Quantitative PCR was carried out using PowerUp SYBR Green Master Mix (ThermoFisher #A25742). Target CT values were normalized to the geometric mean of GAPDH and FL.

**Primary neuron cell culture and transfection**

Cortices from embryonic day (E)19–20 Long-Evans rat embryos were dissected and disassociated, and primary neurons plated at a density of $6x10^5$ cells/mL in 96-well plates, as described previously[100]. At in vitro

day (DIV) 4–5, neurons were transfected with 100 ng of pGW1-mApple[69] to mark cells bodies and 100 ng of an experimental construct (i.e. FUGW-TDP43-EGFP) using Lipofectamine 2000, as before[100]. Following transfection, cells were placed into Neurobasal with B27 supplement (Gibco, Waltham, MA; for all survival experiments). For shRNA knockdown experiments, neurons were transfected with 50 ng of FUGW-EGFP and 200ng shRNA/well. Cells were treated with either SMARTvector Lentiviral Rat control hEF1a-TurboRFP shRNA (Horizon, Lafayette, CO) or SMARTvector Lentiviral Rat Ass1 hEF1a-TurboRFP shRNA (5'-ATTCCAATGAAGCGGTTCT-3') targeting a coding region in the 12th exon of rat ASS1 (V3SR11242-238358331).

**Longitudinal fluorescence microscopy**

Neurons were imaged as described previously[11,12,61,66,67,69] using a Nikon (Tokyo, Japan) Eclipse Ti inverted microscope with PerfectFocus3 and a 20X objective lens. Detection was accomplished with an Andor (Belfast, UK) iXon3 897 EMCCD camera or Andor Zyla4.2 (+) sCMOS camera. A Lambda XL Xenon lamp (Sutter) with 5 mm liquid light guide (Sutter Instrument, Novato, CA) was used to illuminate samples, and custom scripts written in Beanshell for use in μManager controlled all stage movements, shutters, and filters. Custom ImageJ/Fiji macros and Python scripts (https://github.com/barmadaslab/survival-analysis and https://github.com/barmadaslab/measurements; copies archived at https://github.com/elifesciences-publications/survival-analysis and https://github.com/elifesciences-publications/measurements) were used to identify neurons and draw regions of interest (ROIs) based upon size, morphology, and fluorescence intensity. Criteria for marking cell death included rounding of the soma, loss of fluorescence and degeneration of neuritic processes.

**Immunohistochemistry**

Immunostaining was accomplished using the Dako Autostainer Link 48 (Agilent, USA). An anti-pUPF1 antibody (Milipore #07-1016, 1:300) was used with the Dako High pH Target Retrieval Solution (Tris/EDTA, pH 9; Agilent, USA) (20 minutes, 97°C) and the Dako Envision Flex Plus Mouse Link Kit (Agilent, USA) to detect the antibody along with the Dako DAB (Agilent, USA). Whole-slide images were generated by the University of Michigan Digital Pathology group within the Department of Pathology using a Leica Biosystems Aperio AT2 scanner equipped with a 0.75 NA Plan Apo 20x objective; 40x scanning is achieved using a 2x optical magnification changer. Resolution is 0.25μ/pixel for 40x scans. Focus during the scan is maintained using a triangulated focus map built from individual focus points determined in a separate step before scanning is started. Proprietary software is used for image processing during acquisition. Image analysis performed using QuPath software and mean intensity calculated for stained cells.

**ASS1 mRNP pull down assay**

Purification of target-specific mRNPs was adapted from Matia-González et al. (2017)[64]. Briefly, HEK293T cells grown on standard 10cm culture dishes were UV-crosslinked with 254nm UV at 150 mJ/cm2. Cells were harvested by triturating and pelleted at 235g for 10m, discarding the supernatant. The cell pellet was resuspended in 200µl of ice-cold lysis buffer (100 mM Tris–HCl pH 7.4, 500 mM LiCl, 10 mM EDTA, 1% Triton X-100, 5 mM DTT, 20 U/ml Baseline-ZERO DNase (Lucigen #DB0715K), 100 U/ml RNasin (Promega #N2611), and 1x cOmplete EDTA-free Protease Inhibitor Cocktail (Roche #11873580001)) and sonicated with 3 rounds of 20s bursts of level 10 amplitude Q55 Sonicator (Qsonica, Newtown, CT). The lysate was centrifuged at 15,000xg for 10m and the supernatant was collected. Roughly 35ug of protein and 100ul B&W buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl,

0.5 mM EDTA, 32 U/ml RNasin, 10 U/ml Baseline-ZERO DNase, pH 8.0). Biotinylated ASOs were added to lysates to a final concentration of 22.50µM with respect to the volume of Dynabeads MyOne Streptavidin C1 slurry (Thermofisher #65001). The lysate-ASO solution was incubated at 70C for 5m in a thermoblock then removed and allowed to cool to room temperature. In a separate tube, 100µl of streptavidin beads/10cm plate was washed with an equal volume of B&W 3x, using a magnet to separate beads from the supernatant. Washed beads were added to the RT lysate-ASO solution and incubated for 30m at RT on a benchtop rotator. Beads were then collected via magnet and washed 3x with 500µl B&W warmed to 50C. After the final wash, beads were resuspended in 20µl Laemmli buffer for resolution with SDS-PAGE and western blotting.

**SLAM-seq**

On the 14th day in vitro, live lenti-transduced iMNs were metabolically labeled with 400µM 4sU for 45m after which half were collected with Trizol and half were washed, incubated with 10mM unlabeled uridine for 4h, and collected with Trizol. After purification and DNase treatment, RNA was alkylated with a buffered iodoacetemide solution (10 mM iodoacetamide, 50% DMSO, 50 mM sodium phosphate buffer, pH8) at 50C for 15m. The alkylation reaction was stopped with 20mM DTT and the RNA was purified with ethanol precipitation. cDNA libraries were made using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus kit (Illumina, #20040525). Paired end sequencing was carried out on an Illumina NovaSeq (S4) 300 cycle sequencer at the University of Michigan Advanced Genomics Core. Samples were sequenced at an average depth of 75M reads/sample. Trimmed reads were aligned to the hg38 genome using STAR (version 2.7.10b) with optional arguments "--outSAMtype BAM SortedByCoordinate --outSAMattributes NH HI NM MD AS nM jM jI XS --seedSearchLmax 20"[76]. Conversion analysis and kinetic modeling was achieved with GRAND-SLAM software and further analyzed in R with the GrandR package[77].

**Fluorescence in-situ hybridization**

In situ hybridization was performed per manufacturer's recommended protocol (ACDbio Multiplex RNAscope). Fresh-frozen, paraffin-embedded post-mortem tissues mounted on slides were obtained from the University of Michigan Brain Bank (Table 3). Samples were de-paraffinized by immersion in xylene for 5m, then xylene for 5m with agitation, followed by 100% EtOH for 5m and 100% EtOH for 2m with agitation. The slides were dried for 5m at 60C before adding 5-8 drops of $H_2O_2$ (ACDbio #322381) to each and incubating at RT for 10m. Samples were rinsed with $ddH_2O$ 3-5 times before progression to antigen retrieval. Slides were placed in heated $ddH_2O$ (99C) for 10s, 1X retrieval agent (ACDbio #322001) for 15m at 99C, $ddH_2O$ for 15s at RT, 100%EtOH for 3m at RT, then dried at 60C for 5m. Outlines around each sample were drawn with a hydrophobic barrier pen (ACDbio #PN 310018), then 5 drops of Protease Plus (ACDbio #322330) were added prior to placing in a pre-warmed oven (40C) for 15m. Probes for STMN2 (ACDbio #525211-C2) and ASS1 (ACDbio #431291) were resuspended to 1X in probe diluent (ACDbio #300041) before baking in HybEZ II oven (ACDbio #321710/321720) for 2hr at 40C, then rinsing with 1X wash buffer (ACDbio #310091) for 2m at RT. Hybridization was accomplished with Multiplex FL v2 AMP1, AMP2 and AMP3 (ACDbio #323110), each at 40C for 30m, separated by rinses in 1X wash buffer for 2m at RT. For signal development, Multiplex FL v2 HRP-C1 was added to the samples for 15m at 40C. Slides were rinsed twice in 1X wash buffer for 2m at RT before adding 1X TSA Vivid Fluorophore 520 (ACDbio, #323271) for 30m at 40C, rinsing twice in 1X wash buffer for 2m at RT, then baking for 15m at 40C in Multiplex FL vs HRP Blocker (ACDbio, #323110) before rinsing 2x more in 1X wash buffer at RT. The process was repeated for Multiplex FL v2 HRP-C2 and -C3, using 1X TSA Vivid Fluorophore 570 and 1X TSA Vivid Fluorophore 650, respectively. Immunostaining for MAP2 was performed using chicken anti-MAP2 primary antibodies (Novus #NB300-213) diluted 1:3000 in 1% PBS with BSA, ON at 4C. Slides were rinsed three times in 1X wash buffer for 2m at RT before adding secondary antibody (goat anti-chicken Alexa 647, 1:250 in PBS with 1% BSA; Jackson Immunoresearch #103-605-155). Slides were rinsed five times in 1X wash buffer for 2m at RT, and coverslips were mounted using Prolong Gold (without DAPI; ThermoFisher #P36935). Full-thickness Z-stacks were acquired from the anterior horn of each slide using a Nikon (Tokyo, Japan) NSPARC confocal microscope and NIS Elements software. Puncta were quantified from maximal intensity projections of each image in ImageJ/FIJI by drawing regions of interest around individual neurons in the MAP2 channel, and counting puncta in remaining channels using the 'analyze particles' feature. Plots and statistical comparisons for puncta number/cell were created using Superplots and R, respectively[101].

## Acknowledgements

We thank the patients that donated tissue samples to make this work possible. Skin samples from the study participants were collected and de-identified in collaboration with the Michigan Institute for Clinical and Health Research (MICHR, UL1TR000433) through an institutional review board (IRB)-approved protocol

(HUM00028826). We thank Dr. Stephen A. Goutman, Director of the University of Michigan ALS Clinic and Biorepository, and Crystal Pacut from the Program for Neurology Research and Discovery. We also would like to acknowledge Mr. Matthew D. Perkins for his help with postmortem tissue from the University of Michigan Brain Bank, and Kathy Toy for her expertise with immunohistochemistry. Lastly, we are grateful to Francesco Alessandrini and Evangelos Kiskinis from Northwestern University for the generous use of RNA sequencing data.

This work was supported by National Institutes of Health (R01NS097542, R01NS113943, and 1R56NS128110-01 to S.J.B; F31NS115257 to N.G.; F31NS134123 to M.D.; P30AG072931 to the University of Michigan Brain Bank and Alzheimer's Disease Research Center; T32-GM145470 to C.H. and the University of Michigan Cellular and Molecular Biology Graduate Program), the family of Angela Dobson and Lyndon Welch, the A. Alfred Taubman Medical Research Institute, the Danto Family, Ann Arbor Active Against ALS, and the Robert Packard Center for ALS Research. Immunohistochemistry was performed at the Rogel Cancer Center Tissue and Molecular Pathology Shared Resource Laboratory at the University of Michigan (NIH P30 CA04659229).

## Author contributions

N.G. and S.J.B. designed the study. N.G. performed all RNA isolations, bioinformatic analyses, transfections, wide-field fluorescence microscopy, RNA pull-downs and immunoblotting. F.A., E.K., B.Z., J.R., and J.K. provided RNA-seq data. C.A. and Y.B. performed fluorescence in-situ hybridizations and confocal microscopy. E.M.T. maintained iPSCs and differentiated neurons for experiments. X.L. assisted with cell culture, preparation of primary neurons, and RT-PCR. J.W. wrote the code for image acquisition and neuronal survival analysis. C.H. performed roadblock-qPCR. J.K., J.R. and B.Z. provided unique RNA-seq datasets; S.J.B. and N.G. assembled figures, wrote and edited the manuscript.

## Declaration of interests

S.J.B. serves on the advisory board for Neurocures, Inc., Symbiosis, Eikonizo Therapeutics, Ninesquare Therapeutics, the Live Like Lou Foundation, and the Robert Packard Center for ALS Research. S.J.B. has received research funding from Denali Therapeutics, Biogen, Inc., Lysoway Therapeutics, Amylyx Therapeutics, Acelot Therapeutics, Meira GTX, Inc., Prevail Therapeutics, Eikonizo Therapeutics, and Ninesquare Therapeutics.

## Tables

**Table 1: Validation RNA-seq datasets**

| First author, year | Cell type | Control group | NMD-null group |
|---|---|---|---|
| Colombo, 2017[34] | HeLa | siControl | siSMG6/siSMG7 |
| Boehm, 2021[102] | HEK293T | WT; siControl | siSMG6-treated SMG7-/- |
| Zaepfel, 2021 (unpublished) | HEK293T | DMSO | SMG1i |
| Alessandrini, 2021 (unpublished) | iPSC-derived iMNs | siControl | siUPF1 |
| Zaepfel, 2021 (unpublished) | iPSC-derived iMNs | DMSO | SMG1i |
| Gomez, 2021 (this manuscript) | iPSC-derived iMNs | DMSO | SMG1i (11j) |

**Table 2: ALS RNA-seq datasets**

| First author, year | Cell type | Control group | Experimental group | Disease Category |
|---|---|---|---|---|
| Alessandrini, 2021 (unpublished) | iPSC-derived iMNs | isogenic control | *C9ORF72* mutation | *C9ORF72*-ALS/FTD |
| Zaepfel, 2021 (unpublished) | iPSC-derived iMNs | isogenic control | *C9ORF72* mutation | *C9ORF72*-ALS/FTD |
| AnswerALS, 2022[103] | iPSC-derived iMNs | isogenic control | *C9ORF72*-ALS/FTD and sALS | *C9ORF72*-ALS/FTD and sALS |
| Dafinca, 2020[104] | iPSC-derived iMNs | isogenic control | TDP43 (M337V) knock-in | *TARDBP*-ALS/FTD and sALS |

| | | | | |
|---|---|---|---|---|
| Klim, 2019[26] | iPSC-derived iMNs | shControl or EGFP overexpression | TDP43 KD, TDP43 (WT) overexpression, TDP43 (M337V) overexpression | *TARDBP*-ALS/FTD and sALS |
| Liu, 2019[65] | FACS-sorted nuclei from patient cortex | TDP43 positive (no nuclear clearance) | TDP43 negative (nuclear clearance) | *C9ORF72*-ALS/FTD |
| Prudencio, 2015[75] | patient frontal cortex | Healthy control cortex | C9ALS or sALS | *C9ORF72*-ALS/FTD and sALS |
| Tam, 2019[105] | patient cortex | Healthy control cortex | ALS | *C9ORF72*-ALS/FTD and sALS |
| Labadorf, 2015[54] | patient prefrontal cortex | Healthy control cortex | Huntington's disease cortex | Huntington's disease |

**Table 3: Information on post-mortem samples**

| Sample | Age | Sex | PMI | Condition |
|---|---|---|---|---|
| ALS-1 | 64 | Male | 7 | sALS |
| NDC-1 | 68 | Male | 22 | Alzheimer's disease |
| ALS-2 | 68 | Male | 4 | sALS |
| NDC-2 | 91 | Male | 18 | Alzheimer's disease |
| ALS-3 | 63 | Female | 7 | sALS |
| NDC-3 | 75 | Male | 6 | Alzheimer's disease |

NDC = non-ALS disease control

## Figure legends

**Figure 1: The NMDq analytic pipeline.**

**a,** Outline of NMDq raw data processing**. b**, Number of unique annotated and novel isoform models with normal termination codons (NTC), premature termination codons (PTC), and unclassified termination codons from the Colombo et al. (2017)[34] dataset. **c**, Representative gene diagram for *DNAJC17,* and corresponding annotated and novel isoform models. **d**, Transcript abundance, measured as transcripts per million (TPM) for isoforms modeled in (**d**). **e**, gene-level TPM for DNAJC17. **f**, cumulative TPM of isoforms belonging to NTC, PTC, or unclassified categories. Data points in (**d**-**f**) represent biological replicates compared by Wilcoxon rank-sum test ($p < 0.05$, control n = 6, experimental n = 3). **g**, Volcano plot showing differential NMD-class usage (DNU). Data points represent log2 fold differences in PTC to total isoform abundance per gene between experimental and control groups. Abbreviations: STAR: Spliced Transcripts Alignment to a Reference; ORF = open reading frame; ST2: StringTie 2; ISAr: IsoformSwitchAnalyzeR; kd: knockdown.

**Figure 2: NMDq detects PTC burden in diverse cell types and conditions.**

**a**-**f**, Validation of NMDq in immortalized cells. **a**-**c**, **g**-**i**, Top, log2 fold-change density plots comparing relative isoform TPM between NMD-deficient (NMD-) and NMD-intact (NMD+) conditions. Bottom, volcano plots showing the distribution of significance. Dashed black lines indicate the filtering threshold of log2 fold-change ≤ -1 and ≥ 1. Black point corresponds to the commonest PTC-containing BAG1 isoform. All data shown are significant (adjusted p-value ≤ 0.05). **a**, siSMG6- and siSMG7-treated vs siControl-treated HeLa cells (from Colombo et al., 2017)[34]. **b**, siSMG6-treated SMG7-/- vs siControl SMG7+/+ HEK293T cells (from Boehm et al., 2021)[102]. **c**, SMG1 inhibitor (SMG1i)-treated vs DMSO-treated HEK293T cells (from Zaepfel et al., unpublished). **d**, **j**, cumulative, unnormalized TPM of PTC-containing BAG1 (**d**) or SRSF2 (**j**) isoforms in immortalized and iMN datasets, respectively. **e**, **k**, cumulative TPM belonging to PTC+ categories, normalized within the dataset to respective control. Data points represent biological replicates. Black outline represents the cumulative TPM of significantly differentially expressed isoforms (DNU+ genes). **f**, **l**, Euler diagram of the parental genes of differentially expressing PTC+ isoforms. **g**-**i**, Validation of NMDq in iMNs. **g**, siUPF1-treated vs siControl-treated cells (from Alessandrini et al., unpublished). **h**, SMG1i-treated vs DMSO-treated cells (from Zaepfel et al., 2021)[21]. **i**, SMG1i-treated vs DMSO-treated cells (this study).

**Figure 3: NMDq identifies biologically relevant and accessible targets of canonical NMD.**

**a**, First quadrant MA-plot (log-ratio vs mean expression) of immortalized cell validation datasets (from left to right: Colombo et al., 2017[34]: siSMG6- and siSMG7-treated vs siControl-treated HeLa cells; Boehm et al., 2021[102]: siSMG6-treated SMG7-/- vs siControl SMG7+/+ HEK293T cells; and Zaepfel et al., unpublished: SMG1 inhibitor (SMG1i)-treated vs DMSO-treated HEK293T cells). Points represent isoforms. Black points represent isoforms belonging to parental genes with at least one PTC-containing isoform up-regulated in all NMD-deficient conditions. **b**, **e**, the subset of isoforms in (**a**) and supplemental figure 1a that are significantly differentially expressed in all (**b**) or at least two datasets (**e**). Points in red show isoform candidates selected for

follow-up studies. **c**, **f**, Isoform-specific RT-PCR of SMG1i- or DMSO-treated HEK293T cells (**c**) or iMNs (**g**). **d**, isoform-specific roadblock RT-PCR of SMG1i- or DMSO-treated HEK293T cells. DMSO = dimethylsulfoxide; FC = fold change; FF = firefly luciferase.

**Figure 4: ALS models and tissues show no evidence for NMD impairment.**
**a**-**c**, Cumulative TPM belonging to the PTC category, normalized to experiment controls in *C9ORF72*- or sALS-based iMN models (**a**), TDP43-based iMN models (**b**), and post-mortem brain tissue (**c**). Data points represent biological replicates. Error bars represent standard error of the mean. Groups compared by Wilcoxon rank-sum test (*: p<=0.05). **d**, Control-normalized PTC+ fraction of NMD-sensitive targets identified by NMDq from RNA-seq datasets represented in **a**-**c**. Leftmost column represents the geometric mean of targets. **e**, qRT-PCR of NMD-sensitive isoforms in HEK293T overexpressing ALS-associated protein variants. Data points represent biological replicates from 3 experiments compared to the EGFP-DMSO condition by t-test. (*: p<=0.05, **: p<=0.01). **f**-**g**, phospho-UPF1 immunostaining of human spinal cord (**f**) and frontal cortex (**g**). Ghosted points represent cells from at least 2 individuals. Solid points represent means compared by Wilcoxon rank-sum test.

**Figure 5: UPF1 counter-regulates TDP43-associated genes.**
**a**, Volcano plot of TDP43 + GFP vs TDP43 + UPF1 lentivirus-treated iMNs. **a, inset**, top 10 GO terms of significantly upregulated genes in this comparison. **b**, iMN qRT-PCR validation of candidates identified via RNA-seq. **c**, qRT-PCR of *ASS1* and *TARDBP* HEK293T cells transfected with GFP, TDP43, or RNA binding-deficient TDP43. **b**-**c**, data points represent biological replicates for 3 experiments. Indicated comparisons made by Wilcoxon rank-sum test. **d**, Agarose electrophoresis of *ASS1* PCR products purified by biotinylated-ASO pulldown and cDNA reverse transcription. **e**, SDS-PAGE and immunoblotting of mRNP species purified by biotinylated-ASO pulldown, showing successful pulldown of UPF1 and TDP43, but not the negative control (TUB1A).

**Figure 6: ASS1 is an essential factor reflecting TDP43 function in ALS.**
**a**, Single-molecule fluorescence in situ hybridization (smFISH) for *STMN2* and *ASS1* in post-mortem human ALS and neurological disease control (NDC) spinal cord. All tissues were immunostained for MAP2 to aid in identification of motor neurons within the anterior horn. Count of *STMN2* (**b**) and ASS1 (**c**) smFISH puncta in motor neurons from ALS and NDC spinal cord. **p<0.05, Kruskal-Wallis with Dunn's post-hoc test. **d**, Abundance of ASS1 in neurons displaying TDP43 pathology, sorted from the brains of C9ORF72-ALS/FTD patients (adapted from Liu et al., 2019)[65]. **$p_{adj}$=0.002, Wald test with Benjamini–Hochberg correction. **e**, longitudinal microscopy of rodent primary cortical neurons expressing mApple. Neurons are imaged repeatedly over a 10d period, identified by segmentation algorithms (red outlines) and assigned a unique identifier (blue numbers). Cell death is indicated by fragmentation and loss of neurites (green outline) and/or reduction in fluorescence intensity (dark blue outline). **f**, ASS1 immunoreactivity in neurons transfected with non-targeting (NT) or ASS1-directed shRNA, separated into low- and high-expressing groups based on shRNA reporter

levels. Groups are compared by Wilcoxon rank-sum test. **g**, cumulative risk of death for primary neurons transfected with shASS1 or non-targeting shRNA. **h**, cumulative risk of death for primary neurons expressing EGFP or TDP43-EGFP, together with mApple or ASS1. In **g** and **h**, data combined and stratified among 4 separate experiments imaged every 24h for 10d (*$p \leq 0.01$, **$p < 2x10^{-16}$, Cox proportional hazards; n=733-2288 neurons/condition).

**Figure 7: UPF1 restores global 3'UTR length and RNA stability in iMNs overexpressing TDP43.**
**a-c**, mean distal polyA site usage index (DPAU) scatter plots in (**a**) TDP43 vs GFP expressing iMNs, (**b**) TDP43/UPF1 vs GFP expressing iMNs, and (**c**) C9ALS vs Healthy control frontal cortex. **d**, experimental design of SLAM seq in lentivrus treated iMNs. Labeling and sequencing was conducted in biological triplicate. **e**, mismatch rate of each nucleobase as determined by deep RNA sequencing of 4sU-labeled and unlabeled iMNs. **f**, empirical distribution function of RNA half-lives in hours from iMNs expressing indicated lentivirus constructs. **g**, distribution of fold changes in RNA half-lives for indicated comparison (top) among TDP43 targets previously identified in CLIP-seq studies[81,82,106]. **h**-**i**, distribution of fold changes in RNA half-lives (TDP43 vs GFP in **h**, or TDP43/UPF1 vs TDP43 in **i**) for genes that exhibited significant 3'UTR lengthening or non-lengthening changes in TDP43-expressing iMNs compared to GFP-expressing iMNs. Data points in **g**-**i** represent gene-level estimations of RNA half-lives. Distributions are compared by Wilcoxon rank-sum test (****: p < 2 x 10-16). **j**-**k**, representative example of increased distal polyA site usage (**j**) and half-life in hours (**k**) of synaptotagmin 1 (*SYT1*). Error bars represent 95% confidence interval.

## References


1. Charcot, J.-M. (1825-1893) A. du texte. *Leçons Sur Les Maladies Du Système Nerveux Faites à La Salpêtrière. Tome 2 / Par J.-M. Charcot,... ; Recueillies et Publiées Par Bourneville,... [Puis] Charles Féré*. (1872).
2. Goetz, C. G. Amyotrophic lateral sclerosis: Early contributions of Jean-Martin Charcot. 8 (2000).
3. Sabatelli, M., Conte, A. & Zollino, M. Clinical and genetic heterogeneity of amyotrophic lateral sclerosis: Clinical and genetic heterogeneity of ALS. *Clin. Genet.* **83**, 408–416 (2013).
4. Neumann, M. Molecular Neuropathology of TDP-43 Proteinopathies. *Int. J. Mol. Sci.* **10**, 232–246 (2009).
5. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: Disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
6. Ayala, Y. M. *et al.* Human, Drosophila, and C.elegans TDP43: Nucleic Acid Binding Properties and Splicing Regulatory Function. *J. Mol. Biol.* **348**, 575–588 (2005).
7. Ayala, Y. M. *et al.* Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* 8.
8. Barmada, S. J. & Finkbeiner, S. Pathogenic TARDBP Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia: Disease-Associated Pathways. *Rev. Neurosci.* **21**, (2010).
9. Abramzon, Y. A., Fratta, P., Traynor, B. J. & Chia, R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front. Neurosci.* **14**, 42 (2020).
10. Bennion Callister, J. & Pickering-Brown, S. M. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Exp. Neurol.* **262**, 84–90 (2014).
11. Barmada, S. J. *et al.* Cytoplasmic Mislocalization of TDP-43 Is Toxic to Neurons and Enhanced by a Mutation Associated with Familial Amyotrophic Lateral Sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
12. Barmada, S. J. *et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc. Natl. Acad. Sci.* **112**, 7821–7826 (2015).
13. Jackson, K. L. *et al.* Preservation of forelimb function by UPF1 gene therapy in a rat model of TDP-43-induced motor paralysis. *Gene Ther.* **22**, 20–28 (2015).
14. Wils, H. *et al.* TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3858–3863 (2010).

15. Liachko, N. F., Guthrie, C. R. & Kraemer, B. C. Phosphorylation Promotes Neurotoxicity in a Caenorhabditis elegans Model of TDP-43 Proteinopathy. *J. Neurosci.* **30**, 16208–16219 (2010).

16. Kabashi, E. *et al.* Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum. Mol. Genet.* **19**, 671–683 (2010).

17. Godfrey, R. K. *et al.* Modelling TDP-43 proteinopathy in Drosophila uncovers shared and neuron-specific targets across ALS and FTD relevant circuits. *Acta Neuropathol. Commun.* **11**, 168 (2023).

18. Johnson, B. S., McCaffery, J. M., Lindquist, S. & Gitler, A. D. A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proc. Natl. Acad. Sci.* **105**, 6439–6444 (2008).

19. Ju, S. *et al.* A Yeast Model of FUS/TLS-Dependent Cytotoxicity. *PLoS Biol.* **9**, e1001052 (2011).

20. Sun, Y., Eshov, A., Zhou, J., Isiktas, A. U. & Guo, J. U. C9orf72 arginine-rich dipeptide repeats inhibit UPF1-mediated RNA decay via translational repression. *Nat. Commun.* **11**, 3354 (2020).

21. Zaepfel, B. L. *et al.* UPF1 reduces C9orf72 HRE-induced neurotoxicity in the absence of nonsense-mediated decay dysfunction. *Cell Rep.* **34**, 108925 (2021).

22. Ortega, J. A. *et al.* Nucleocytoplasmic Proteomic Analysis Uncovers eRF1 and Nonsense-Mediated Decay as Modifiers of ALS/FTD C9orf72 Toxicity. *Neuron* **106**, 90-107.e13 (2020).

23. Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Amp Dev.* **21**, 1833–3856 (2007).

24. Kim, Y. K. & Maquat, L. E. UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. *RNA* **25**, 407–422 (2019).

25. Ling, J. P., Pletnikova, O., Troncoso, J. C. & Wong, P. C. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* **349**, 650–655 (2015).

26. Klim, J. R. *et al.* ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat. Neurosci.* **22**, 167–179 (2019).

27. Melamed, Z. *et al.* Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat. Neurosci.* **22**, 180–190 (2019).

28. Brown, A.-L. *et al.* TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. *Nature* **603**, 131–137 (2022).

29. Ma, X. R. *et al.* TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A. *Nature* **603**, 124–130 (2022).

30. Seddighi, S. *et al. Mis-Spliced Transcripts Generate* de Novo *Proteins in TDP-43-Related ALS/FTD*. http://biorxiv.org/lookup/doi/10.1101/2023.01.23.525149 (2023) doi:10.1101/2023.01.23.525149.

31. Tank, E. M. *et al.* Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat. Commun.* **9**, 2845 (2018).

32. Paillusson, A. A GFP-based reporter system to monitor nonsense-mediated mRNA decay. *Nucleic Acids Res.* **33**, e54–e54 (2005).

33. Gerbracht, J. V., Boehm, V. & Gehring, N. H. Plasmid transfection influences the readout of nonsense-mediated mRNA decay reporter assays in human cells. *Sci. Rep.* **7**, 10616 (2017).

34. Colombo, M., Karousis, E. D., Bourquin, J., Bruggmann, R. & Mühlemann, O. Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA* **23**, 189–201 (2017).

35. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

36. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).

37. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).

38. Vitting-Seerup, K. & Sandelin, A. IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinforma. Oxf. Engl.* **35**, 4469–4471 (2019).

39. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).

40. Green, R. E. *et al.* Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* **19**, i118–i121 (2003).

41. Kovalak, C., Donovan, S., Bicknell, A. A., Metkar, M. & Moore, M. J. Deep sequencing of pre-translational mRNPs reveals hidden flux through evolutionarily conserved alternative splicing nonsense-mediated decay pathways. *Genome Biol.* **22**, 132 (2021).

42. Olivieri, J. E. *et al.* RNA splicing programs define tissue compartments and cell types at single-cell resolution. *eLife* **10**, e70692 (2021).

43. Huang, L. *et al.* RNA Homeostasis Governed by Cell Type-Specific and Branched Feedback Loops Acting on NMD. *Mol. Cell* **43**, 950–961 (2011).

44. Medghalchi, S. M. *et al.* Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* **10**, 99–105 (2001).

45. Leclair, N. K. *et al.* Poison Exon Splicing Regulates a Coordinated Network of SR Protein Expression during Differentiation and Tumorigenesis. *Mol. Cell* **80**, 648-665.e9 (2020).

46. Gopalsamy, A. *et al.* Identification of pyrimidine derivatives as hSMG-1 inhibitors. *Bioorg. Med. Chem. Lett.* **22**, 6636–6641 (2012).

47. Karam, R. *et al.* The unfolded protein response is shaped by the NMD pathway. *EMBO Rep.* **16**, 599–609 (2015).

48. Karousis, E. D., Gypas, F., Zavolan, M. & Mühlemann, O. Nanopore sequencing reveals endogenous NMD-targeted isoforms in human cells. *Genome Biol.* **22**, 223 (2021).

49. Watson, M. J., Park, Y. & Thoreen, C. C. Roadblock-qPCR: a simple and inexpensive strategy for targeted measurements of mRNA stability. *RNA* **27**, 335–342 (2021).

50. Slomovic, S., Laufer, D., Geiger, D. & Schuster, G. Polyadenylation of ribosomal RNA in human cells. *Nucleic Acids Res.* **34**, 2966–2975 (2006).

51. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).

52. Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72**, 257–268 (2011).

53. Janssens, J. *et al.* Overexpression of ALS-Associated p.M337V Human TDP-43 in Mice Worsens Disease Features Compared to Wild-type Human TDP-43 Mice. *Mol. Neurobiol.* **48**, 22–35 (2013).

54. Labadorf, A. *et al.* RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. *PloS One* **10**, e0143563 (2015).

55. Yang, L. *et al.* Transcriptomic Landscape of von Economo Neurons in Human Anterior Cingulate Cortex Revealed by Microdissected-Cell RNA Sequencing. *Cereb. Cortex* **29**, 838–851 (2019).

56. Kim, E.-J. *et al.* Selective frontoinsular von Economo neuron and fork cell loss in early behavioral variant frontotemporal dementia. *Cereb. Cortex N. Y. N 1991* **22**, 251–259 (2012).

57. Gami-Patel, P. *et al.* Von Economo neurons are part of a larger neuronal population that are selectively vulnerable in C9orf72 frontotemporal dementia. *Neuropathol. Appl. Neurobiol.* **45**, 671–680 (2019).

58. Moya, M. V. *et al.* Unique molecular features and cellular responses differentiate two populations of motor cortical layer 5b neurons in a preclinical model of ALS. *Cell Rep.* **38**, 110556 (2022).

59. Lee, K. M., Chand, K. K., Hammond, L. A., Lavidis, N. A. & Noakes, P. G. Functional decline at the aging neuromuscular junction is associated with altered laminin-α4 expression. *Aging* **9**, 880–899 (2017).

60. Engel, K., Höhne, W. & Häberle, J. Mutations and polymorphisms in the human argininosuccinate synthetase (ASS1) gene. *Hum. Mutat.* **30**, 300–307 (2009).

61. Flores, B. N. *et al.* An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep.* **27**, 1133 (2019).

62. Bhardwaj, A., Myers, M. P., Buratti, E. & Baralle, F. E. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* **41**, 5062–5074 (2013).

63. Tseng, S.-H., Cheng, C.-Y., Huang, M.-Z., Chung, M.-Y. & Su, T.-S. Modulation of formation of the 3'-end of the human argininosuccinate synthetase mRNA by GT-repeat polymorphism. *Int. J. Biochem. Mol. Biol.* **4**, 179–190 (2013).

64. Matia-González, A. M., Jabre, I. & Gerber, A. P. Biochemical approach for isolation of polyadenylated RNAs with bound proteins from yeast. *STAR Protoc.* **2**, 100929 (2021).

65. Liu, E. Y. *et al.* Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. *Cell Rep.* **27**, 1409-1421.e6 (2019).

66. Archbold, H. C. *et al.* TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci. Rep.* **8**, 4606 (2018).

67. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *eLife* **7**, e35977 (2018).

68. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R. & Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810 (2004).

69. Barmada, S. J. *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat. Chem. Biol.* **10**, 677–685 (2014).

70. Rot, G. *et al.* High-Resolution RNA Maps Suggest Common Principles of Splicing and Polyadenylation

Regulation by TDP-43. *Cell Rep.* **19**, 1056–1067 (2017).

71. Ayala, Y. M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop: TDP-43 autoregulation of its transcript levels. *EMBO J.* **30**, 277–288 (2011).
72. Xia, Z. *et al.* Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat. Commun.* **5**, 5274 (2014).
73. Hogg, J. R. & Goff, S. P. Upf1 senses 3′UTR length to potentiate mRNA decay. *Cell* **143**, 379–389 (2010).
74. Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M. & Mühlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA N. Y. N* **17**, 2108–2118 (2011).
75. Prudencio, M. *et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175–1182 (2015).
76. Alalam, H., Zepeda-Martínez, J. A. & Sunnerhagen, P. Global SLAM-Seq for accurate mRNA decay determination and identification of NMD targets. 29.
77. Jürges, C., Dölken, L. & Erhard, F. Dissecting newly transcribed and old RNA using GRAND-SLAM. *Bioinforma. Oxf. Engl.* **34**, i218–i226 (2018).
78. Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
79. Tan, S. *et al.* Post-transcriptional regulation of ERBB2 by miR26a/b and HuR confers resistance to tamoxifen in estrogen receptor-positive breast cancer cells. *J. Biol. Chem.* **292**, 13551–13564 (2017).
80. Tushev, G. *et al.* Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **98**, 495-511.e6 (2018).
81. Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
82. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).
83. Hurt, J. A., Robertson, A. D. & Burge, C. B. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res.* **23**, 1636–1650 (2013).
84. Imamachi, N., Salam, K. A., Suzuki, Y. & Akimitsu, N. A GC-rich sequence feature in the 3′ UTR directs UPF1-dependent mRNA decay in mammalian cells. *Genome Res.* **27**, 407–418 (2017).

85. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **21**, 708–718 (2007).

86. Ratnadiwakara, M. & Änkö, M.-L. mRNA Stability Assay Using transcription inhibition by Actinomycin D in Mouse Pluripotent Stem Cells. *Bio-Protoc.* **8**, e3072 (2018).

87. Held, A. *et al.* iPSC motor neurons, but not other derived cell types, capture gene expression changes in postmortem sporadic ALS motor neurons. *Cell Rep.* **42**, 113046 (2023).

88. Kamelgarn, M. *et al.* ALS mutations of FUS suppress protein translation and disrupt the regulation of nonsense-mediated decay. *Proc. Natl. Acad. Sci.* **115**, E11904–E11913 (2018).

89. Yoshitoshi-Uebayashi, E. Y. *et al.* Modelling urea-cycle disorder citrullinemia type 1 with disease-specific iPSCs. *Biochem. Biophys. Res. Commun.* **486**, 613–619 (2017).

90. Chandler, R. J. *et al.* Liver-directed adeno-associated virus serotype 8 gene transfer rescues a lethal murine model of citrullinemia type 1. *Gene Ther.* **20**, 1188–1191 (2013).

91. Zünd, D., Gruber, A. R., Zavolan, M. & Mühlemann, O. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3′ UTRs. *Nat. Struct. Mol. Biol.* **20**, 936–943 (2013).

92. Kurosaki, T. *et al.* A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev.* **28**, 1900–1916 (2014).

93. Kishor, A. *et al.* Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. *Nucleic Acids Res.* **48**, 7468–7482 (2020).

94. Sato, H. & Singer, R. H. Cellular variability of nonsense-mediated mRNA decay. *Nat. Commun.* **12**, 7203 (2021).

95. Zheng, D. *et al.* Cellular stress alters 3’UTR landscape through alternative polyadenylation and isoform-specific degradation. *Nat. Commun.* **9**, 2268 (2018).

96. Chen, M. *et al.* 3’ UTR lengthening as a novel mechanism in regulating cellular senescence. *Genome Res.* **28**, 285–294 (2018).

97. Paulsen, M. T. *et al.* Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods* **67**, 45–54 (2014).

98. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

99. Chua, J. P. *et al.* Myotubularin-related phosphatase 5 is a critical determinant of autophagy in neurons. *Curr. Biol. CB* S0960-9822(22)00671–6 (2022) doi:10.1016/j.cub.2022.04.053.

100. Weskamp, K., Safren, N., Miguez, R. & Barmada, S. Monitoring Neuronal Survival via Longitudinal Fluorescence Microscopy. *J. Vis. Exp.* 59036 (2019) doi:10.3791/59036.

101. Goedhart, J. SuperPlotsOfData-a web app for the transparent display and quantitative comparison of continuous data from different conditions. *Mol. Biol. Cell* **32**, 470–474 (2021).

102. Boehm, V. *et al.* SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. *Nat. Commun.* **12**, 3965 (2021).

103. Baxi, E. G. *et al.* Answer ALS, a large-scale resource for sporadic and familial ALS combining clinical and multi-omics data from induced pluripotent cell lines. *Nat. Neurosci.* **25**, 226–237 (2022).

104. Dafinca, R. *et al.* Impairment of Mitochondrial Calcium Buffering Links Mutations in C9ORF72 and TARDBP in iPS-Derived Motor Neurons from Patients with ALS/FTD. *Stem Cell Rep.* **14**, 892–908 (2020).

105. Tam, O. H. *et al.* Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *Cell Rep.* **29**, 1164-1177.e5 (2019).

106. Hallegger, M. *et al.* TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell* **184**, 4680-4696.e22 (2021).

**Supplemental material**

# Counter-regulation of RNA stability by UPF1 and TDP43


## Author list

Nicolas Gomez[1], Caroline Hsieh[1], Xingli Li[2], Megan Dykstra[3], Jacob Waksmacki[2], Christopher Altheim[2], Yoel Bechar[2], Joseph Klim[4], Benjamin Zaepful[5], Jeffrey Rothstein[6], Elizabeth EM Tank[2], Sami J Barmada[1,2]

## Affiliations

[1]Graduate Program in Cell and Molecular Biology, University of Michigan, Ann Arbor, MI

[2]Department of Neurology, University of Michigan, Ann Arbor, MI

[3]Neuroscience Graduate Program, University of Michigan, Ann Arbor, MI

[4]Aquinnah Pharmaceuticals, Cambridge, MA

[5]Department of Neurology, University of California, San Diego, CA

[6]Department of Neurology, Johns Hopkins University, Baltimore, MD

*To whom correspondence should be addressed

Sami Barmada

University of Michigan

Department of Neurology

109 Zina Pitcher Place, 4019 BSRB

Ann Arbor, MI 48109

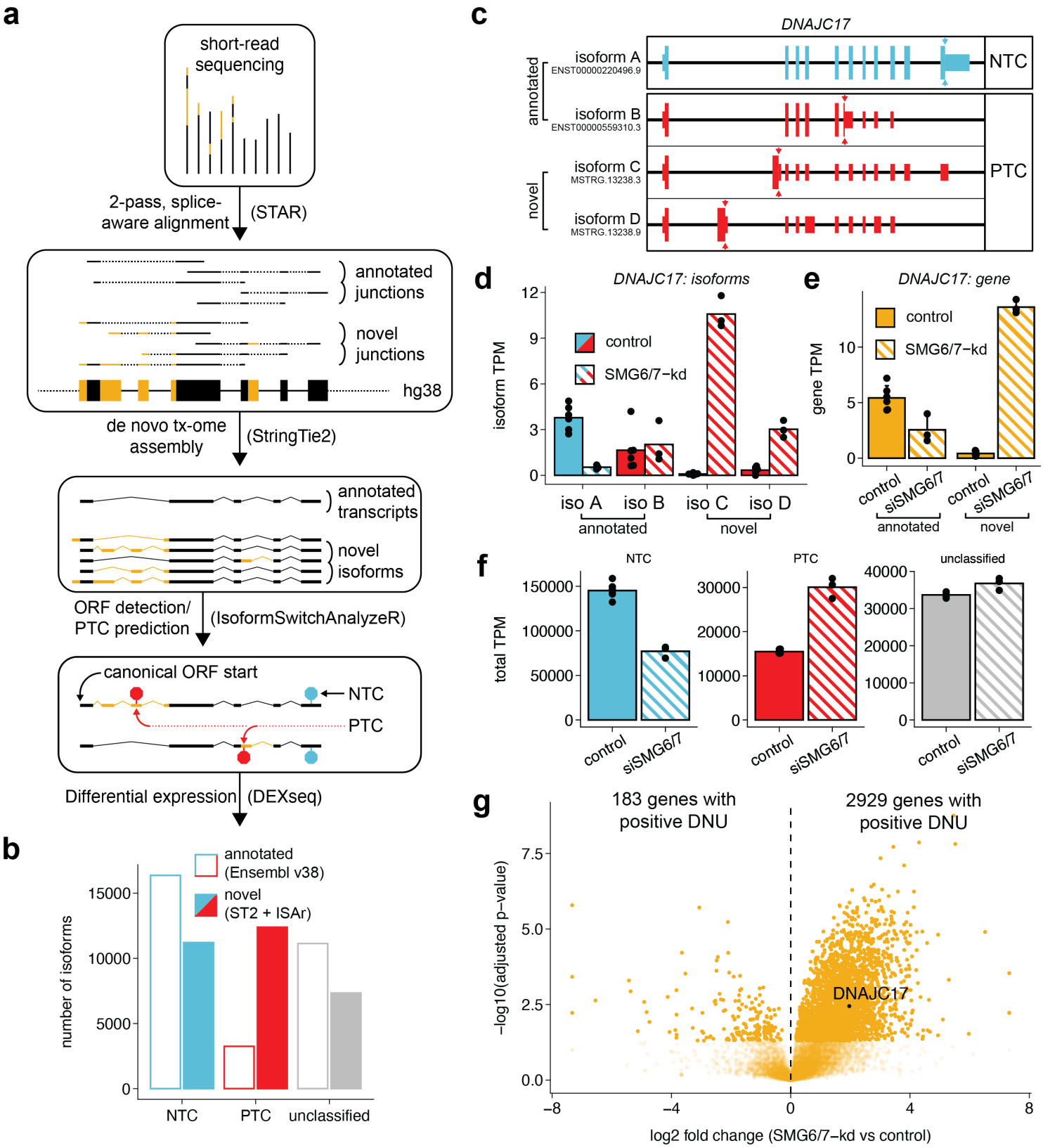
**Supplemental Figure 1: NMDq identifies assayable, biologically relevant targets of canonical NMD (continued from Figure 3).**

**a**, First quadrant MA-plot (log-ratio vs mean expression) of iMN validation datasets (from left to right: Alessandrini, 2021: siUPF1-treated vs siControl-treated iMNs; Zaepfel, 2021: SMG1i-treated vs DMSO-treated iMNs; and Gomez, 2021: SMG1i-treated vs DMSO-treated iMNs). Points represent isoforms. Black points represent isoforms belonging to parental genes with at least one PTC-containing isoform up-regulated in two NMD-deficient conditions. **b**, SDS-PAGE and immunoblotting of HEK293T cells treated with DMSO or SMG1i and stained for GAPDH and SAT1. Data represents three biological replicates compared by paired Wilcoxon signed-rank test ($p = 0.25$). **c**, representative RNA-seq reads from DMSO- or 11j-treated iMNs of poison or putative poison exons at the *SRSF2*, *SAT1*, and *BAG1* loci.

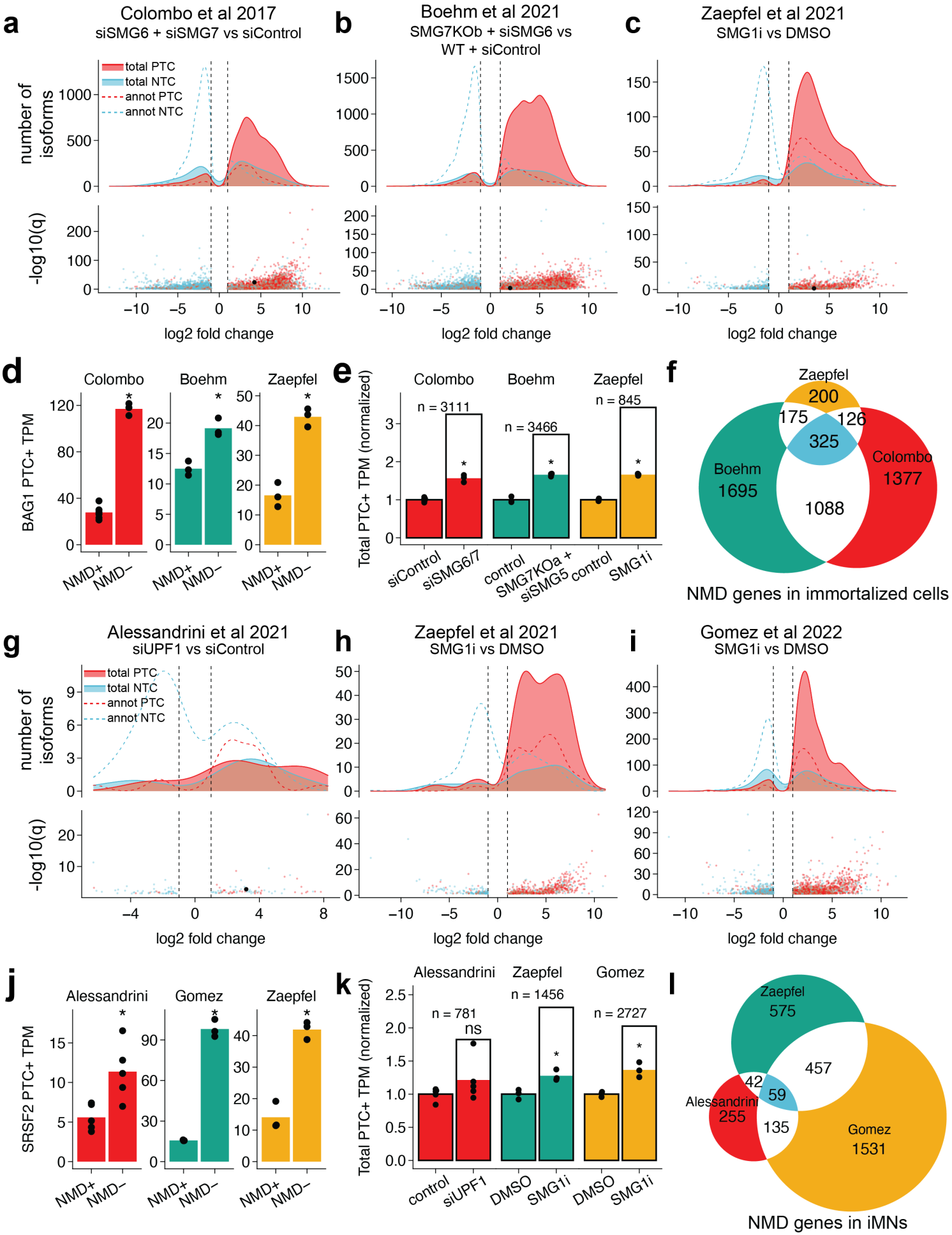
**Supplemental Figure 2: NMD is not impaired in ALS models or tissues (continued from Figure 4).**

Cumulative TPM belonging to the PTC category in lentivirus-treated iMNs, normalized to GFP/DMSO treated condition. Data points represent biological replicates. Error bars represent standard error of the mean. Groups compared by Wilcoxon rank-sum test (*: $p<=0.05$).

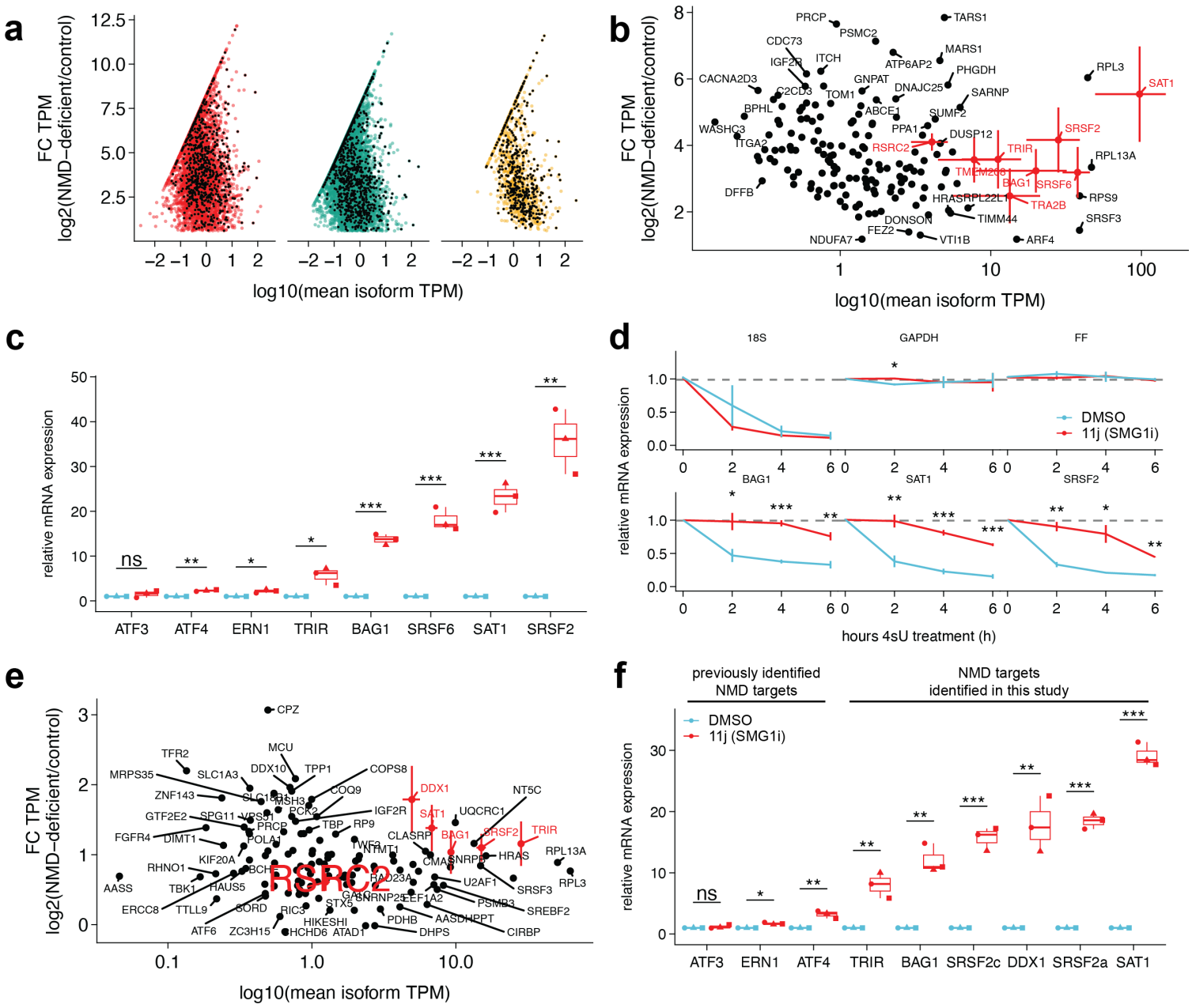
# Figure 1


a
short-read sequencing
2-pass, splice-aware alignment
(STAR)
annotated junctions
novel junctions
hg38
de novo tx-ome assembly
(StringTie2)
annotated transcripts
novel isoforms
ORF detection/ PTC prediction
(IsoformSwitchAnalyzeR)
canonical ORF start
NTC
PTC
Differential expression
(DEXseq)
b
annotated (Ensembl v38)
novel (ST2 + ISAr)
number of isoforms
15000
10000
5000
0
NTC
PTC
unclassified
c
DNAJC17
annotated
novel
isoform A
ENST00000220496.9
isoform B
ENST00000559310.3
isoform C
MSTRG.13238.3
isoform D
MSTRG.13238.9
NTC
PTC
d
DNAJC17: isoforms
isoform TPM
control
SMG6/7−kd
iso A
iso B
iso C
iso D
annotated
novel
e
DNAJC17: gene
gene TPM
control
SMG6/7−kd
control
siSMG6/7
annotated
novel
f
NTC
PTC
unclassified
total TPM
control
siSMG6/7
g
183 genes with positive DNU
2929 genes with positive DNU
−log10(adjusted p−value)
DNAJC17
log2 fold change (SMG6/7−kd vs control)

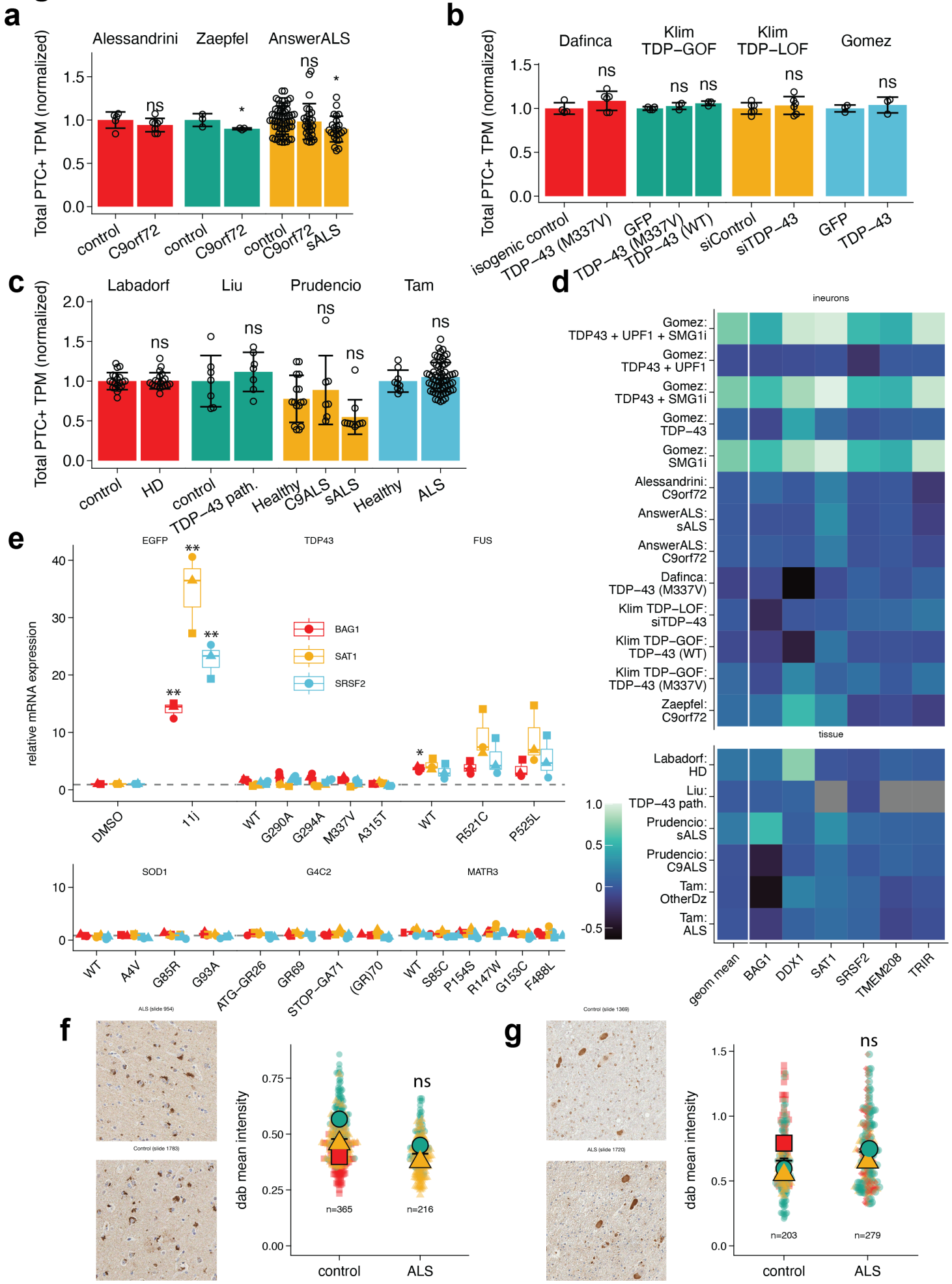
# Figure 2


a
Colombo et al 2017
siSMG6 + siSMG7 vs siControl
b
Boehm et al 2021
SMG7KOb + siSMG6 vs WT + siControl
c
Zaepfel et al 2021
SMG1i vs DMSO
total PTC
total NTC
annot PTC
annot NTC
number of isoforms
-log10(q)
log2 fold change
d
Colombo
Boehm
Zaepfel
BAG1 PTC+ TPM
NMD+
NMD−
e
Total PTC+ TPM (normalized)
n = 3111
n = 3466
n = 845
siControl
siSMG6/7
control
SMG7KOa + siSMG5
SMG1i
f
Zaepfel
200
175
126
325
Boehm
1695
Colombo
1377
1088
NMD genes in immortalized cells
g
Alessandrini et al 2021
siUPF1 vs siControl
h
Zaepfel et al 2021
SMG1i vs DMSO
i
Gomez et al 2022
SMG1i vs DMSO
j
Alessandrini
Gomez
Zaepfel
SRSF2 PTC+ TPM
k
Alessandrini
Zaepfel
Gomez
Total PTC+ TPM (normalized)
n = 781
ns
n = 1456
n = 2727
control
siUPF1
DMSO
SMG1i
l
Zaepfel
575
457
42
59
Alessandrini
255
135
Gomez
1531
NMD genes in iMNs

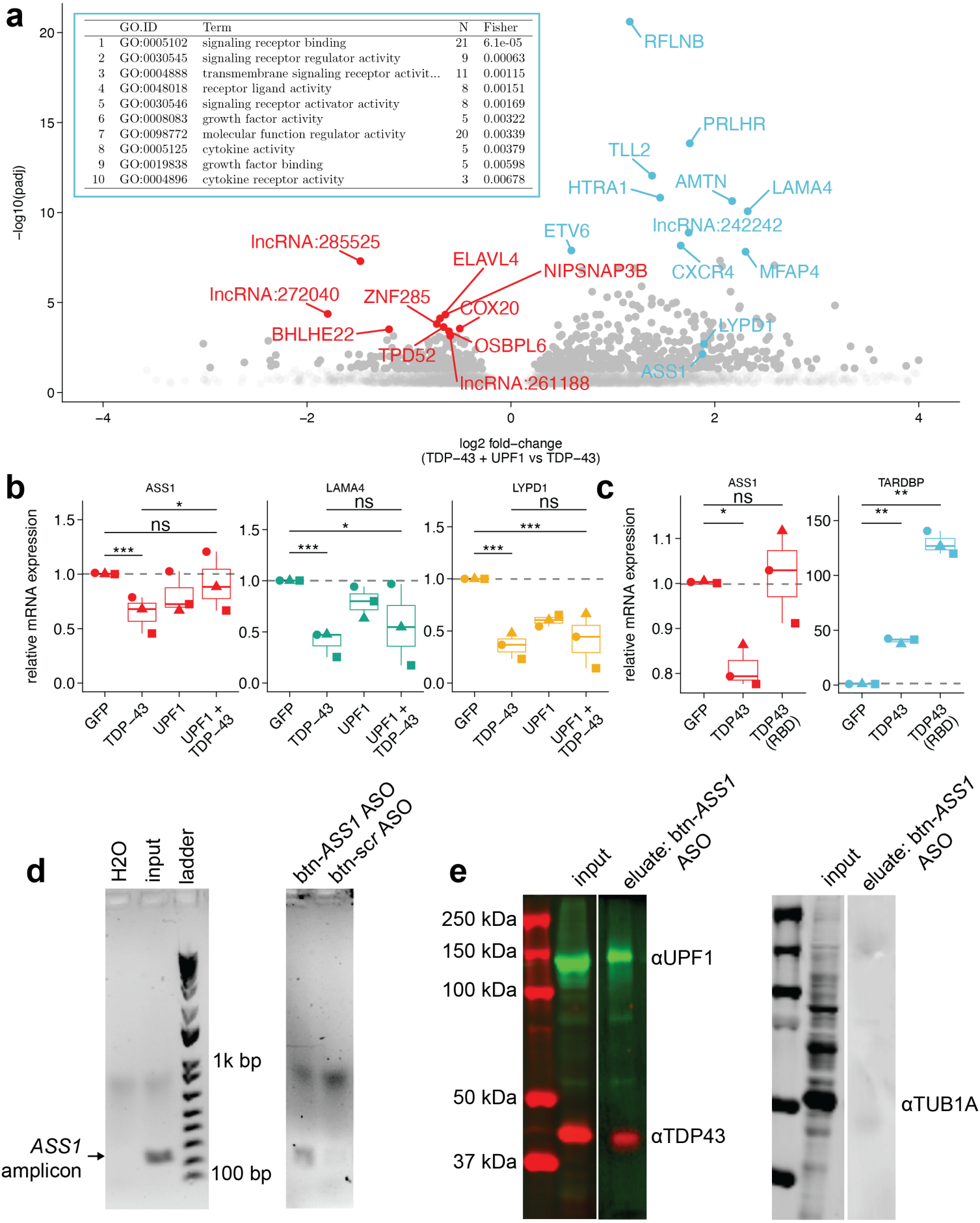
# Figure 3


a
FC TPM
log2(NMD−deficient/control)
log10(mean isoform TPM)
b
FC TPM
log2(NMD−deficient/control)
log10(mean isoform TPM)
PRCP
PSMC2
TARS1
CDC73
MARS1
IGF2R
ITCH
ATP6AP2
PHGDH
CACNA2D3
GNPAT
DNAJC25
SARNP
C2CD3
TOM1
ABCE1
BPHL
SUMF2
WASHC3
PPA1
DUSP12
ITGA2
RSRC2
TRIR
TMEM208
BAG1
SRSF6
TRA2B
HRASRPL22L1
DFFB
DONSON
TIMM44
FEZ2
NDUFA7
VTI1B
ARF4
RPL3
SAT1
SRSF2
RPL13A
RPS9
SRSF3
c
relative mRNA expression
ATF3
ATF4
ERN1
TRIR
BAG1
SRSF6
SAT1
SRSF2
ns
d
18S
GAPDH
FF
DMSO
11j (SMG1i)
BAG1
SAT1
SRSF2
relative mRNA expression
hours 4sU treatment (h)
e
FC TPM
log2(NMD−deficient/control)
log10(mean isoform TPM)
CPZ
MCU
TFR2
MRPS35
SLC1A3
DDX10
TPP1
COPS8
ZNF143
COQ9
DDX1
NT5C
GTF2E2
SPG11
IGF2R
SAT1
UQCRC1
FGFR4
DIMT1
TBP
RP9
CLASRP
BAG1
SRSF2
TRIR
RPL13A
KIF20A
HRAS
RHNO1
AASS
TBK1
HAUS5
U2AF1
RPL3
SRSF3
ERCC8
TTLL9
SORD
PSMB3
SREBF2
ATF6
ZC3H15
HIKESHI
HCHD6
ATAD1
PDHB
AASDHPPT
DHPS
CIRBP
RSRC2
f
previously identified NMD targets
NMD targets identified in this study
DMSO
11j (SMG1i)
relative mRNA expression
ATF3
ERN1
ATF4
TRIR
BAG1
SRSF2c
DDX1
SRSF2a
SAT1
ns

# Figure 4

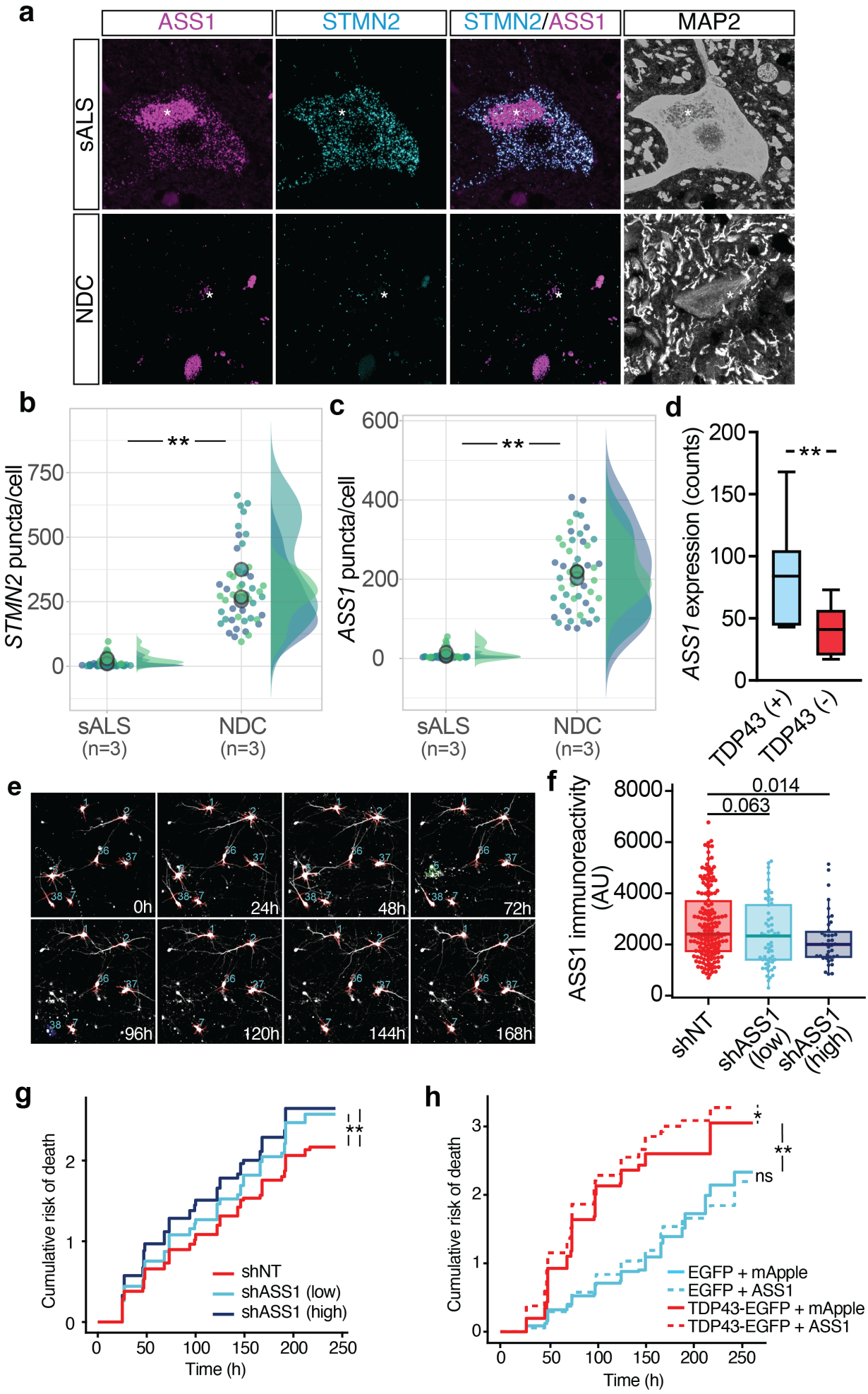

a
Alessandrini
Zaepfel
AnswerALS
Total PTC+ TPM (normalized)
ns
*
control
C9orf72
sALS
b
Dafinca
Klim TDP−GOF
Klim TDP−LOF
Gomez
isogenic control
TDP−43 (M337V)
GFP
TDP−43 (WT)
siControl
siTDP−43
c
Labadorf
Liu
Prudencio
Tam
HD
TDP−43 path.
Healthy
C9ALS
ALS
d
ineurons
Gomez: TDP43 + UPF1 + SMG1i
Gomez: TDP43 + UPF1
Gomez: TDP43 + SMG1i
Gomez: TDP−43
Gomez: SMG1i
Alessandrini: C9orf72
AnswerALS: sALS
AnswerALS: C9orf72
Dafinca: TDP−43 (M337V)
Klim TDP−LOF: siTDP−43
Klim TDP−GOF: TDP−43 (WT)
Klim TDP−GOF: TDP−43 (M337V)
Zaepfel: C9orf72
tissue
Labadorf: HD
Liu: TDP−43 path.
Prudencio: sALS
Prudencio: C9ALS
Tam: OtherDz
Tam: ALS
geom mean
BAG1
DDX1
SAT1
SRSF2
TMEM208
TRIR
e
EGFP
TDP43
FUS
relative mRNA expression
BAG1
SAT1
SRSF2
DMSO
11j
WT
G290A
G294A
M337V
A315T
R521C
P525L
SOD1
G4C2
MATR3
A4V
G85R
G93A
ATG−GR26
GR69
STOP−GA71
(GR)70
S85C
P154S
R147W
G153C
F488L
f
ALS (slide 954)
Control (slide 1783)
dab mean intensity
n=365
n=216
control
ALS
g
Control (slide 1369)
ALS (slide 1720)
n=203
n=279

# Figure 5

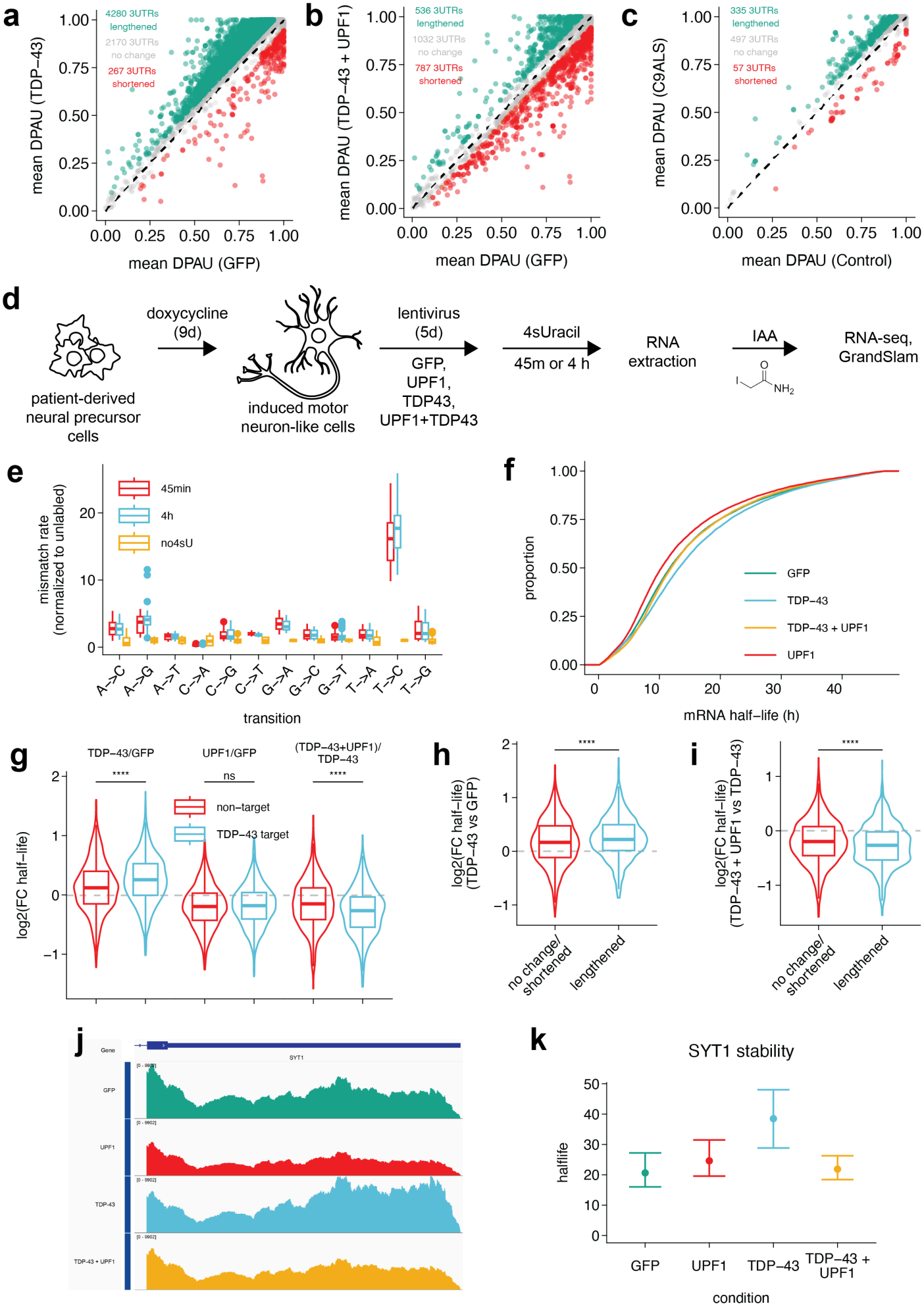



| | GO.ID | Term | N | Fisher |
|---|---|---|---|---|
| 1 | GO:0005102 | signaling receptor binding | 21 | 6.1e-05 |
| 2 | GO:0030545 | signaling receptor regulator activity | 9 | 0.00063 |
| 3 | GO:0004888 | transmembrane signaling receptor activit... | 11 | 0.00115 |
| 4 | GO:0048018 | receptor ligand activity | 8 | 0.00151 |
| 5 | GO:0030546 | signaling receptor activator activity | 8 | 0.00169 |
| 6 | GO:0008083 | growth factor activity | 5 | 0.00322 |
| 7 | GO:0098772 | molecular function regulator activity | 20 | 0.00339 |
| 8 | GO:0005125 | cytokine activity | 5 | 0.00379 |
| 9 | GO:0019838 | growth factor binding | 5 | 0.00598 |
| 10 | GO:0004896 | cytokine receptor activity | 3 | 0.00678 |

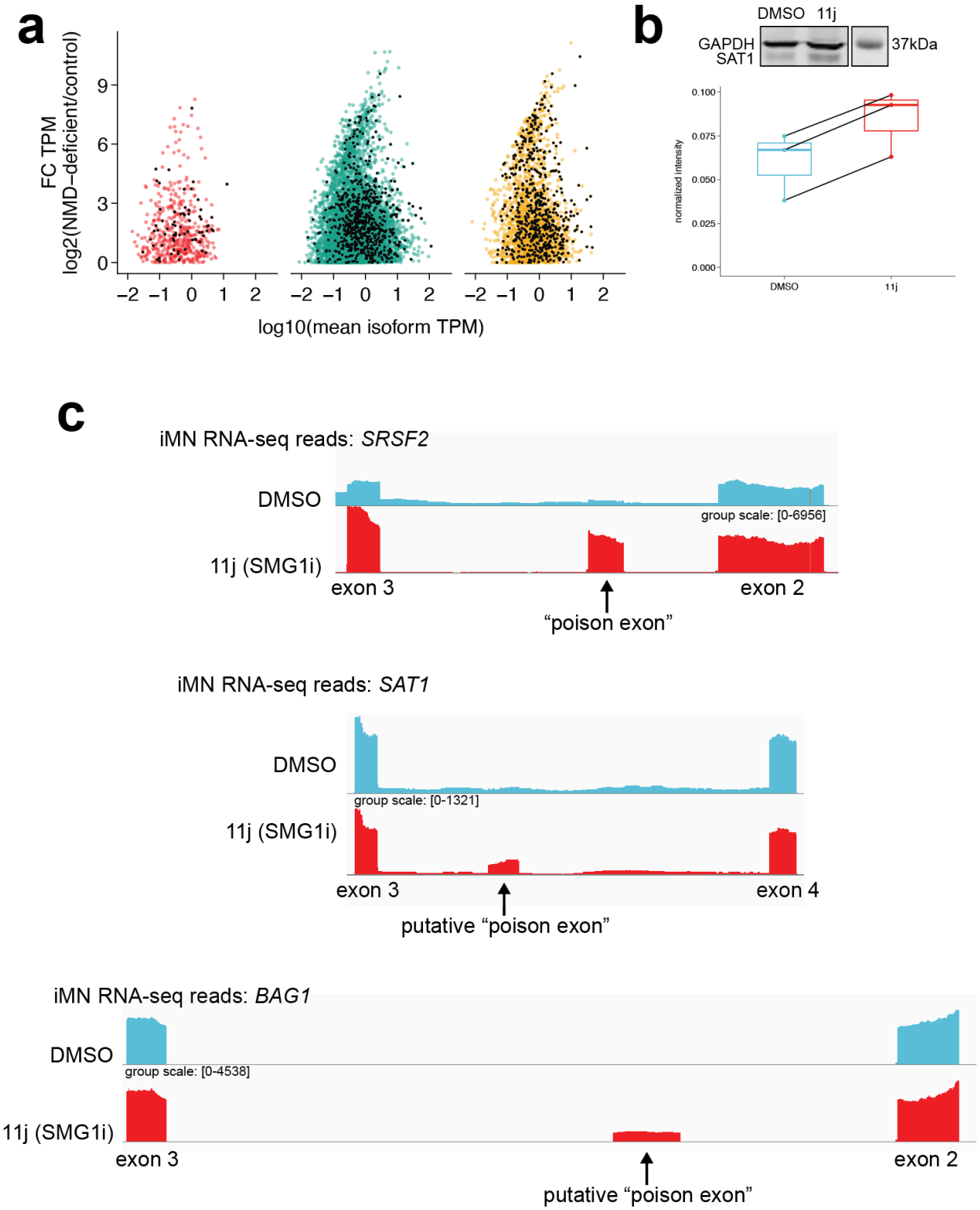
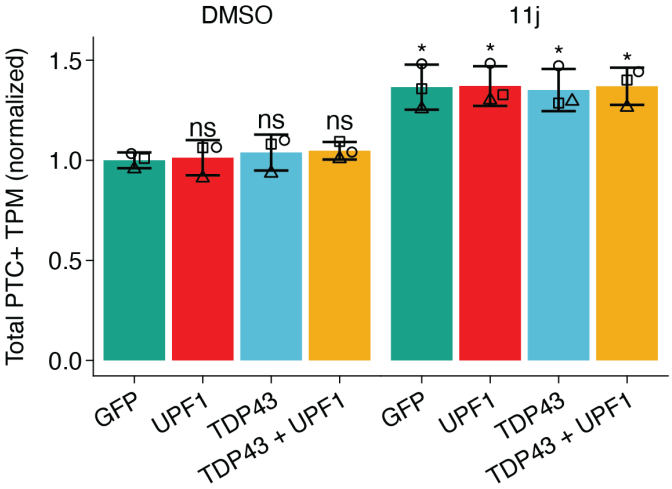
# Figure 6


a
ASS1
STMN2
STMN2/ASS1
MAP2
sALS
NDC
b
STMN2 puncta/cell
750
500
250
0
**
sALS
(n=3)
NDC
(n=3)
c
ASS1 puncta/cell
600
400
200
0
**
sALS
(n=3)
NDC
(n=3)
d
ASS1 expression (counts)
200
150
100
50
0
**
TDP43 (+)
TDP43 (-)
e
0h
24h
48h
72h
96h
120h
144h
168h
f
ASS1 immunoreactivity (AU)
8000
6000
4000
2000
0
0.014
0.063
shNT
shASS1 (low)
shASS1 (high)
g
Cumulative risk of death
2
1
0
**
shNT
shASS1 (low)
shASS1 (high)
0
50
100
150
200
250
Time (h)
h
Cumulative risk of death
3
2
1
0
*
**
ns
EGFP + mApple
EGFP + ASS1
TDP43-EGFP + mApple
TDP43-EGFP + ASS1
0
50
100
150
200
250
Time (h)

# Figure 7


a
4280 3UTRs lengthened
2170 3UTRs no change
267 3UTRs shortened
mean DPAU (TDP−43)
mean DPAU (GFP)
b
536 3UTRs lengthened
1032 3UTRs no change
787 3UTRs shortened
mean DPAU (TDP−43 + UPF1)
mean DPAU (GFP)
c
335 3UTRs lengthened
497 3UTRs no change
57 3UTRs shortened
mean DPAU (C9ALS)
mean DPAU (Control)
d
patient-derived neural precursor cells
doxycycline (9d)
induced motor neuron-like cells
lentivirus (5d)
GFP, UPF1, TDP43, UPF1+TDP43
4sUracil
45m or 4 h
RNA extraction
IAA
RNA-seq, GrandSlam
e
45min
4h
no4sU
mismatch rate (normalized to unlabled)
transition
f
GFP
TDP−43
TDP−43 + UPF1
UPF1
proportion
mRNA half−life (h)
g
TDP−43/GFP
UPF1/GFP
(TDP−43+UPF1)/ TDP−43
non−target
TDP−43 target
log2(FC half−life)
h
log2(FC half−life) (TDP−43 vs GFP)
no change/ shortened
lengthened
i
log2(FC half−life) (TDP−43 + UPF1 vs TDP−43)
no change/ shortened
lengthened
j
Gene
SYT1
GFP
UPF1
TDP-43
TDP-43 + UPF1
k
SYT1 stability
halflife
GFP
UPF1
TDP−43
TDP−43 + UPF1
condition

# Supplemental figure 1



# Supplemental figure 2