**Intersection Between RNA Methylation and TDP43-Mediated Toxicity in ALS**

by

Michael McMillan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor Sami Barmada, Chair
Professor Anthony Antonellis
Professor Hank Paulson
Professor Nils Walter

Michael P. McMillan

mcmm@umich.edu

ORCID iD:  0000-0002-2067-9941

## Dedication

This dissertation is dedicated to my uncle, Kevin Gosnell, and all those battling ALS in hopes that this research offers a new prospect towards ending this disease.

**Acknowledgements**

First off, I'd like thank Sami for his mentoring and support over the last five years. Coming into grad school I had limited knowledge of the neurodegenerative field or molecular biology to that end and I appreciate Sami taking me in as a member of the lab, trusting me to try my hand at a fun—although complex—project. I enjoyed how he always pushed the envelope to try new techniques and pushed me to learn all aspects of science: the theory, the technical aspects, the communication, and the people. I'd also like to thank my thesis committee—Drs. Tony Antonellis, Hank Paulson, and Nils Walter for their support throughout the years. Committee meetings were never the anxiety-inducing events I'd heard of, and their willingness to share their knowledge and experience during these meetings was very valuable to me as a scientist.

I would also like to acknowledge Dr. Elizabeth Tank, who was an instrumental part of my PhD experience. I'd like to think I've come a long way since my first day of my rotation, and she is a major reason for that. Having Sami and Elizabeth in the lab is like getting two PI's for the price of one, something I really was fortunate to have. She also recruited me to her softball team, which has been a staple of my summers here and through which I've met some great people.

Continuing with lab, I have to thank all the past and present Barmada lab members I've worked with. I was lucky to join at a time where I could learn from experienced members like Brittany, Nate, Kate, Ahmed, Roberto, and Xingli. I'd like to thank Nico, who joined around the same time as me, for always making me think

through experiments and for his valuable bioinformatic support. Lastly, I've enjoyed working with the new wave of Barmada lab members: Genesis, Michael, Megan, Josephine, and Caroline, and hope that I have been able to provide some useful mentorship over the last couple of years. The Barmada lab has been and will continue to be a great environment to work in because we are so close-knit, and I hope future members will have the same memorable experiences as I did. I'd also like to thank members of the Paulson and Todd labs for their expertise, and generosity lending reagents when something was inevitably backordered. Finally, I want to thank the Cellular and Molecular Biology program, the Neurology department, and associated staff for their help throughout the years.

Last but not least, I'd like to thank my family and friends for keeping me going. Grad school can be difficult at times and with a global pandemic thrown on top, it was these people who kept me focused so I could get to where I am today.

## Table of Contents

## List of Tables

## List of Figures

## Abstract

Amyotrophic Lateral Sclerosis (ALS) is a devastating neurodegenerative disease characterized by progressive loss of motor neurons that results in muscle loss, paralysis, and respiratory failure. ALS diagnosis is complicated by its heterogeneous biochemical, genetic, and clinical features. In addition, >90% of ALS cases arise without previous family history or genetic cause of disease, while only <10% result from an inherited mutation in a known ALS causing gene. However, the vast majority of ALS cases present with dysregulation of a key RNA binding protein TDP43. A primarily nuclear protein, TDP43 is mislocalized to the cytoplasm in ALS, accumulating in cytoplasmic inclusions. TDP43 is an essential protein involved with many levels of RNA processing such as alternative splicing and RNA transport, largely influencing RNA stability. Previous studies uncovered widespread RNA destabilization in models of ALS, implicating dysregulation of RNA stability in disease.

RNA stability can be modified by many of the over 150 known RNA modifications. The most common internal RNA modification, methylation of RNA at the N6 position of adenosine (m6A), significantly contributes to RNA stability as well as RNA transport and translation. m6A is co-transcriptionally deposited across many species of RNA at specific motifs by methyltransferases (“writers”) and removed by demethylases (“erasers”). m6A alters the fate of RNA depending on which RNA binding protein (“reader”) binds the methylated RNA, thereby influencing the transport, translation, and stability of that RNA.

In this dissertation, I determine whether and how m6A-modified RNA influences TDP43 RNA substrate recognition and processing. Chapter 1 begins with a review of TDP43, focusing on its structure and functions in RNA processing then characterizing its role in ALS. Chapter 1 also introduces m6A, methods to measure m6A, the function of pathway components, and ties to neurodegenerative disease. Chapter 2 explores the hypothesis that TDP43 recognizes m6A methylated RNA, identifying methylation sites on TDP43 substrates and increased methylation levels in patient tissue. Additionally, removal of an m6A site in the *TARDBP* transcript reduced TDP43s ability to autoregulate and modulation of m6A components can alter TDP43-associated toxicity. Of these, knockout of the m6A reader YTHDF2 rescued toxicity in rodent and human neuronal models of ALS. YTHDF2 is upregulated in patient spinal cord offering a novel therapeutic target. Chapter 3 summarizes key findings from this research and discusses open questions surrounding the overlap between TDP43 and m6A. Finally, Appendix A describes an improved method for CRISPR sgRNA cloning and creating multiplexed vectors as a method to create genome edited cell lines.

# Chapter 1: Introduction

## 1.1 Overview of Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD)

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease resulting from the death of upper and lower motor neurons in the cerebral cortex, brainstem and spinal cord[1]. There are an estimated 6,000 ALS cases per year and an incidence of 2 in 100,000 people[2]. ALS onset ranges from 40-70 years old, with an average age of onset of 55 years[2]. Manifestation of disease is highly clinically variable, but the most common initial symptom is gradual muscle weakness, which soon develops into paralysis. Despite the variance in the disease, muscle weakness and paralysis are observed in all cases. Paralysis eventually spreads to other limbs and then to other muscles such as those involved in the respiratory system. Respiratory failure results in death on average of 2-5 years following disease onset[2]. ALS arises sporadically (sALS) in ~90% of cases, and through familial inheritance (fALS), in ~10% of cases. ~80% of individuals with fALS have a known genetic cause whereas fewer than 20% of individuals with sALS can be explained by genetics[3]. Additionally, as many as 50% of ALS patients can have some sort of cognitive impairment indicating an overlap with frontotemporal dementia (FTD). FTD is the second most common form of dementia in people under 65 years old behind Alzheimer's, and patients with FTD experience issues with personality, behavior, and language due to atrophy [4]. Despite the tremendous diversity in ALS genetics and clinical presentation, 95% of ALS cases

are unified by the presence of cytoplasmic TDP43[5]. This chapter will provide basic background information on TDP43 and its related functions in DNA/RNA binding, RNA metabolism, evidence in disease, and introduce m6A RNA modifications and the proteins involved in this pathway.

## 1.2 TDP43 Structure, Localization, and Function

### *1.2.1 TDP43 structure*

Transactive response DNA binding protein of 43 kDa (TDP43) is the protein encoded by the gene *TARDBP*. TDP43 was first identified because of its ability to bind to bind pyrimidine rich motifs in the TAR region and repress expression of HIV-1 *in vivo* in 1995[6]. TDP43 shared sequence homology with other heterogeneous ribonucleoprotein (hnRNP) proteins such as hnRNPC and hnRNPA2B1, and therefore was characterized as a member of the hnRNP family of proteins. A key characteristic shared amongst this family of proteins is their ability to bind RNA through a highly conserved RNA recognition motif (RRM) domain[6]. Further exploration into the TDP43 protein itself identified that it is 414 amino acids (aa), encoding several domains throughout the span of the protein. The N-terminus, which is aa 1-102, is important for TDP43 homodimerization that may act to regulate protein function[7,8]. TDP43 is primarily a nuclear protein and contains a nuclear localization signal (NLS) from aa 82-98 which regulates nuclear import of TDP43[9,10]. The domains that shared similarity with other hnRNP proteins, RNA recognition motif (RRM1, aa 104-176) and RRM2 (aa 192-262), are the main functional domains where they bind strongly to RNA and DNA in order to regulate several cellular processes such as mRNA processing, export, and stability[11–13]. The RRMs themselves have slightly different activity for binding RNA, where RRM1 has

a higher affinity for binding RNA substrates, and RRM2 appears to improve specificity[11,14,15]. The C-terminus (aa 277-414) contains a prion like domain (aa 345-366) and a glycine rich domain (aa 366-414) which are critical to protein-protein interactions, and splicing activity[9,12,16–18]. Most of the ALS-causing mutations within TDP43 are located in the C-terminal region, suggesting it may facilitate proper TDP43 function and interactions, although not yet proven[19].

### *1.2.2 TDP43 nucleic acid binding properties*

TDP43, like other hnRNP family member proteins, contains RRMs that bind to RNA/DNA and regulate various steps in RNA processing. Investigating proteins that recognize splicing regulatory elements in the *CFTR* gene, Buratti and Baralle found TDP43 recognized and bound to a UG rich element on the 3' splice site of exon 9[13]. Via electromobility shift assays (EMSAs) they show that TDP43 binds to UG repeats, with increased binding as the repeat size increases. Additionally, deletion of one or both of the RRMs decreased RNA binding capabilities. Cross-linking immunoprecipitation sequencing (CLIP-seq) of mouse brains using a TDP43 antibody to isolate TDP43 RNA targets found that TDP43 again bound UG rich sequences, mostly within intronic regions and the 3' UTR of transcripts[20]. Additional studies confirmed that TDP43 strongly binds UG RNA and TG DNA probes, and mutations that disrupt normal RRM interaction with RNA limit RNA binding and reduce toxicity[11,13,15,21]. These data indicate that RNA binding is essential for TDP43-dependent neurodegeneration.

### *1.2.3 TDP43 localization*

TDP43 mostly resides within the nucleus but early data demonstrates TDP43 can shuttle back and forth from the nucleus to the cytoplasm[9]. Movement to and from the nucleus appears to be mediated by RNA binding as expression of full length TDP43, disease associated mutations, or TDP43 lacking an NLS, but not in the RRMs lead to cytoplasmic TDP43[9,22–25]. Overexpression of TDP43 in yeast, worm, flies, mice, and primary rat neurons is sufficient to produce cytoplasmic TDP43 that results in neurodegeneration and ALS like phenotypes[26–32]. Similarly, expression of TDP43 lacking an NLS in mice results in ALS like phenotypes paired with decreased endogenous TDP43[33]. Thus, TDP43 RNA binding ability is critical to its localization, and cytoplasmic TDP43 causes ALS phenotypes in model systems.

### *1.2.4 Alternative splicing*

TDP43 has several functions relating to RNA homeostasis, and one of TDP43's main responsibilities is alternative splicing. Alternative splicing is when exons are either included or excluded from a mature transcript which creates different RNA and protein isoforms from the same gene. This process offers cells a dynamic way to regulate gene expression by influencing RNA stability effectively without requiring going back to the level of transcription. Nearly all mRNAs are alternatively spliced and the brain has the most alternatively spliced mRNA of all organs[34–36]. TDP43 binds UG rich regions on roughly one third of all RNAs, facilitating alternative splicing by repressing inclusion of unannotated exons, or "cryptic exons" [20,37–40]. TDP43 interacts with several splicing factors, and loss of TDP43 results in transcriptome wide changes in alternative splicing[20,37]. Often times, alternative splicing produces mRNAs that contain a premature stop codon, signaling these RNAs for degradation via the nonsense mediated decay

(NMD) pathway[41]. Inclusion of cryptic exons can have severe detrimental effects on RNA stability and cell viability, and many of the alternatively spliced TDP43 regulated RNAs are specific to neurons[42]. The cell type specific changes in alternative splicing suggest that neurons are more susceptible to cell death after disruptions to normal alternative splicing, resulting in neurodegeneration[43]. Loss of other RNA binding proteins that function in alternative splicing by repressing cryptic exons also results in neurodegeneration, implying that correct alternative splicing is critical for neuronal survival[39,40,44,45].

### *1.2.5 TDP43 autoregulation*

TDP43 binds thousands of RNA transcripts, including its own transcript as a mechanism to control the levels of TDP43 in the cell. As an essential protein, TDP43 levels are tightly regulated so that there is not too little or too much protein. TDP43 knock-out in mice is embryonic lethal, and partial or conditional knockout results in neurodegeneration with behavior problems that are similar to that in ALS[38,46–48]. Conversely, overexpression of TDP43 across several different model systems ranging from primary neurons to primates leads to neurodegeneration as well. Maintenance of TDP43 levels is critical to cell survival, and through a negative feedback loop TDP43 is able to do this. TDP43 binds to a specific region in its 3' UTR called the TDP43 binding region (TBR) which triggers alternative splicing[20,37,49,50]. This alternative splicing effectively destabilizes the RNA transcript leading to increased degradation which, in turn, results in less protein being made[20,50,51].

### *1.2.6 TDP43 in ALS*

In ALS, TDP43 is mislocalized to the cytoplasm where it forms inclusions. There are additional TDP43 pathology phenotypes seen in ALS such as nuclear depletion or skein-like formations strengthening the conclusion that disruption of TDP43 proteostasis is central to ALS pathology[52]. As mentioned previously, even subtle overexpression or knock down of TDP43 is toxic to neurons[32], indicating that deviations in TDP43 abundance seen in post-mortem human brain tissue are likely integral to disease pathogenesis. In accordance with this, mutations within the 3' UTR have been shown to cause ALS, potentially because TDP43 is unable to bind its own transcript[19]. However, most pathogenic mutations in *TARDBP* are located in the glycine rich domain[19,32] where it is believed these mutations mislocalize TDP43 to the cytoplasm and interfere with autoregulation such that more TDP43 is produced[32,53,54].

Recent work has identified TDP43 depletion as a cause for cryptic exon inclusions in one gene linked to ALS, Unc-13 homolog A (UNC13A), and connected another, stathmin-2 (STMN2). STMN2 is a regulator of microtubule stability that in the background of TDP43 knockdown is significantly decreased in ALS iPSC derived human motor neurons (hMNs) and SH-SY5Y cells. Also, STMN2 is decreased in ALS spinal cord motor neurons, signifying that STMN2 can act as a biomarker of disease. In normal conditions, TDP43 binds to *SMTN2* and correctly splices out introns to make a fully functional RNA. In disease and with TDP43 depletion, a cryptic exon is included between exon 1 and 2, which results in early truncation of the RNA and subsequent degradation[55,56]. Shortly after STMN2 was discovered, another TDP43 regulated target was identified that acts in a similar manner. UNC13A has previously been identified by Genome wide association studies (GWAS) to be an ALS-FTD risk gene and further

exploration identified a cryptic exon in between exons 21 and 22 and intron retention between 31 and 32[57,58]. Similarly as with STMN2, TDP43 knockdown resulted in decreased expression of UNC13A and in patient tissue the cryptic exon was significantly expressed[56,57]. These are only two examples of what may be a more prominent effect of TDP43 depletion or dysregulation affecting proper splicing resulting in disease.

### *1.2.7 TDP43 & RNA stability: stress granules, miRNA, and RNA transport*

There is emerging evidence that TDP43 is crucial for RNA stability, and that RNA stability may contribute to ALS pathogenesis[59]. In experiments performed by our lab, overexpressing TDP43 in induced pluripotent stem cells (iPSCs) destabilized oxidative phosphorylation pathway RNAs and protein encoding ribosomal RNA[59]. Similarly, RNA stability was measured in iPSCs containing mutations in the C9ORF72 gene, which is the most common mutation found in familial ALS. Mutations in C9ORF72 result in a hexanucleotide repeat expansion of GGGGCC, located in the first intron of C9ORF72. The repeat expansion results in loss of one alternatively spliced *C9ORF72* isoform and nuclear RNA foci[60,61]. The significance of these two pathways in ALS pathology needs further exploration, as does the impact of TDP43 on transcripts that make up those pathways. However, TDP43 plays an integral role in other aspects of RNA stability such as modulating stress granule formation, initiating miRNA processing, and transporting RNAs for local translation.

Stress granules are cytoplasmic structures formed in response to various forms of cellular stress that contain mRNAs, RNA binding proteins, and stalled translation initiation complexes[62–64]. TDP43 localizes to stress granules under osmotic and

oxidative stress, and has been found to colocalize with stress granule components in brain tissue[65–68]. Stress granule dynamics are somewhat influenced by TDP43 because in instances of TDP43 knockdown, stress granule formation is delayed[65,68,69]. ALS causing *TARDBP* mutations increase the rate of stress granule formation, leading to larger granules than with normal TDP43 expression[66,67]. Stress granule dynamics in neurons exhibit different properties in neurons than other cell types, and loss of TDP43 in aged neurons decreased stress granule assembly[70]. These data suggest that TDP43 responds to certain stress conditions, and further investigation is needed to uncover if stress granules are harmful or helpful to disease pathogenesis.

Another method for TDP43 to modulate RNA stability is by promoting biogenesis of micro RNAs (miRNAs). miRNAs are small (20-25 nucleotides) non-coding RNAs that bind within mRNAs to prevent them from being translated or trigger their degradation. miRNA is formed after its precursor, pri-miRNA, forms a hairpin loop after transcription[71,72]. The enzyme Drosha recognizes the hairpin and cleaves it from the transcript and the transcript is exported into the cytoplasm where another enzyme Dicer cuts away the loop of the hairpin to yield two complimentary strands of RNA[73–76]. The two strands break away from each other and join with the RNA-induced silencing complex (RISC) which guides the miRNA to its target. TDP43 interjects at several points in this process, via direct association with pri-miRNA, pre-miRNA, Drosha, and Dicer[77]. These associations help TDP43 regulate the formation of miRNAs essential for neuronal development, activity and survival[77–81]. In ALS cerebrospinal fluid, there is evidence of altered TDP43-regulatd miRNA expression compared to healthy controls[82,83]. In humans carrying a *TARDBP* mutation, neurons had reduced levels of

miRNA 9 (miR-9) and its precursor pri-miR-9-2[80]. These findings imply that TDP43 is important for regulating miRNAs, and if dysregulated like in ALS, could lead to widespread issues of RNA stability given miRNAs can destabilize many mRNA targets[84].

In addition to its other functions in RNA stability, TDP43 regulates transport of RNA, which is especially important for neurons where RNA undergoes local translation at distal regions of the cell[85,86]. In *Drosophila* and mouse models, TDP43 forms cytoplasmic messenger ribonucleoprotein (mRNP) granules that undergo bidirectional microtubule dependent transport to deliver mRNAs to distal parts of neurons. These actions are impaired if TDP43 has ALS-causing mutations, and there is increased anterograde transport in these models[85,86]. TDP43 associates with other RNA binding proteins to regulate retrograde and anterograde transport of mRNP granules. In mouse models, TDP43 cooperates with RNA binding protein Fragile X messenger ribonucleoprotein (FMRP) to modulate anterograde transport and regulate retrograde transport of mRNPs with Staufen1 (STAU1) in neuronal dendrites[87]. TDP43 appears to be critical for axonal elongation, as depletion significantly decrease axon length due to downregulation of cytoskeleton proteins[88]. Further analysis revealed a decreased translation rate within mouse axons and a reduced number of mitochondria. While some of these findings are representative of TDP43 dysfunction, its specific effect within axons points to the importance of TDP43 for neuronal transport and local translation of certain transcripts[88].

ALS causing mutations or TDP43 depletion appear to disrupt stress granule formation, miRNA regulation, and mRNP transport which are critical cellular functions.

Taken together, these data indicate that TDP43 greatly influences RNA stability through regulation of stress granules, miRNA biogenesis, and mRNP transport, which are critical for neuronal survival.

### 1.3 N6-methyladenosine RNA modification

RNA homeostasis can be modulated by many different RNA modifications, however the consequences of most of these modifications are unknown[89]. One of them, N6-methyladenosine methylation (m6A), influences the stability of RNA[90–92]. m6A RNA methylation is the result of an additional methyl group added to the 6$^{th}$ position nitrogen and is the most common internal RNA modification of the over 150 known RNA modifications to date. m6A modifications can be found on nearly every species of RNA including but not limited to messenger RNA (mRNA), ribosomal RNA (rRNA), and long noncoding RNA (lncRNA) and it is estimated that 0.1-0.4% of all adenosines (A) are m6A methylated. Levels of methylation vary across cell and tissue types, but is particularly enriched in the brain and nervous system[93].

Deposition of m6A on RNA begins with methyltransferases (“writers”, described in 1.3.1) that transfer a methyl group from a S-adenosylmethionine donor to a selected adenosine on the RNA. Specifically, this modification occurs within a highly conserved DRACH motif (where D= A, G or U; R= A or G; and H= A, C or U), although only a small fraction of DRACH motifs are actually methylated, for reasons that are still unknown. Methylation takes place at defined regions within the cell, as m6A writers and demethylases (“erasers”, described in 1.3.2) localize to nuclear speckles, membraneless organelles that host factors involved in transcription and splicing[94–98] within the nucleus. Methyltransferases recognize RNA as its being transcribed by RNA

Polymerase II (Pol II) and add m6A co-transcriptionally[99–102]. Importantly, methylation only occurs within the nucleus, and no mechanisms to date are capable of m6A methylating cytoplasmic RNAs.

Despite being discovered in the early 1970s, technological limitations prevented further investigation into m6A and its impact on RNA. It wasn't until the creation of whole transcriptome m6A detection methods were more intricate details of methylation brought to the surface. Using MeRIP-seq (Methylation RNA immunoprecipitation sequencing) or m6A-seq, the methylation profile of the transcriptome showed enrichment of m6A sites near the stop codon and 3' UTR of a transcript[93,103]. Importantly, these studies confirmed the motif for m6A to be within a DRACH context, confirming the sequence bias of the m6A writers.

One drawback to the MeRIP-seq method is relatively low resolution because it relies on the size of the RNA fragments, resulting in a peak range of 50-200 nt. This can complicate detection of m6A sites because there can be multiple DRACH motifs within that region and only one may be methylated, or there can be a cluster of m6A sites. To improve the level of detection, a new method was created called miCLIP-seq (m6A individual nucleotide resolution cross linking immunoprecipitation sequencing)[104]. miCLIP relied upon the finding that when cross linked to RNA, certain m6A antibodies introduced mutations or truncations in the cDNA when reverse transcribed[104]. This increase in sensitivity allowed single-nucleotide resolution of m6A sites compared to the much broader range of MeRIP-seq enabling future experiments to be performed at specific m6A sites. However, a significant caveat of antibody-based approaches for m6A detection is these antibodies also recognize another form of RNA methylation, N6

2’*O*-dimethyladenosine (m6Am). Unlike m6A, m6Am is found at the 5’ end of RNA, usually at the first nucleotide after the m7G cap, and is abundant throughout the transcriptome[105].

In order to address some of the limitations with antibody based approaches, a further improvement on identifying m6A sites came about with a new technique called DART-seq (Deamination adjacent to RNA targets RNA sequencing)[106]. This method relies upon APOBEC1, a cytosine deaminase, fused to the m6A binding domain YTH. The consistency of the DRACH motif where every methylated A is followed by a C allows APOBEC1-YTH to locate m6A sites and deaminate the following C, changing it to a T that can then be tracked by comparison with reference sequences. As with previous approaches, DART-seq found a similar methylation profile as with MeRIP-seq and miCLIP, but with DART-seq only 1 out of 3,431 RNAs had an overlap of m6A and m6Am sites[106]. Thus, DART-seq appears to isolate primarily only m6A sites and is a valuable tool for evaluating m6A across the transcriptome.

An important discovery in the m6A field came when demethylases (erasers, described in 1.3.2) were found to remove methyl marks from adenosine, implying that m6A is a dynamic and controllable modification. FTO and ALKBH5, are two erasers initially believed to function by reverting m6A back to A. In *in vitro* studies, FTO subtly decreases methylation levels, and *in vivo* knockdown or overexpression appeared to change levels of m6A as well[98]. ALKBH5 was identified as another m6A eraser shortly after from a biochemical screen of demethylase proteins that exhibit m6A demethylase activity[95]. ALKBH5 knockout slightly increased m6A levels whereas overexpression slightly decreased levels, indicating it has specific activity towards m6A. However, FTO

was later discovered to act on m6Am, exhibiting almost 100 times stronger catalytic activity against the similar modification. This preference was not seen with ALKBH5, suggesting it still regulates m6A removal or prevents addition[107].

After an RNA is methylated, m6A RNA binding proteins (readers, described below) recognize the methylated RNA and bring the RNA to the next RNA processing step. Many reader proteins contain a YTH domain, responsible for directly binding to m6A. However, indirect readers such as HNRNPC can also bind methylated RNAs due to m6A altering the secondary structure of the RNA allowing certain proteins to bind via an "m6A switch"[108]. Also, other RNA binding domains such as RNA recognition motifs (RRMs) have m6A binding ability suggesting that even though proteins don't contain YTH domains, they can bind methylated RNA[92,103].

### *1.3.1 m6A methyltransferases ("writers")*

#### *1.3.1.1 METTL3*

m6A modifications are added via an m6A methyltransferase complex made up of several proteins that function in concert with each other. m6A-modified RNA was first identified in in the 1970s from fractions of poly(A) RNA, but was difficult to evaluate further due to lack of technologically capable of identifying m6A on RNAs[109,110]. Development of an *in vitro* methylation assay confirmed the existence of m6A sites *in vivo* within a specific sequence context, providing the first evidence that m6A is deposited at specific motifs of GGACU [111,112]. Utilizing this assay, several nuclear fractions were isolated by column chromatography and their respective m6A activity was tested in search of identifying components responsible[90]. In doing so, a 70 kDa fraction with *S*-adenosylmethionine (SAM) was isolated, and following further

purification and sequencing, was determined to be an m6A methyltransferase, termed MT-70[94]. Later renamed to METTL3, MT-70 is part of a larger family of putative SAM-dependent methyltransferases that are highly conserved in mammals[113]. It became evident that METTL3 was the key component to m6A methylation, as genetically depleting METTL3 in yeast, plants, and mammalian cells led to a significant decrease or elimination of m6A on polyA RNA[114–116].

#### *1.3.1.2 METTL14*

In a proteomics study to capture proteins that interact with METTL3, a homologous methyltransferase was found in the analysis and dubbed METTL14[91,97,117,118]. METTL14 is required for m6A formation, and purified METTL14 has methyltransferase activity[117]. Structurally similar to other methyltransferases and with a SAM binding domain, there were notions that METTL14 could be adding m6A to different parts on RNA or to other forms of RNA within the transcriptome. This theory was soon debunked after a series of crystallization experiments identified METTL3 as the only SAM binding protein in the METTL3-METTL14 complex, and determined critical residues within the catalytic domain of METTL3 for m6A deposition[119–121]. Importantly, these studies also determined that while METTL14 has a SAM binding domain, it is not catalytically active and therefore cannot transfer m6A to RNA. METTL14 forms a complex with METTL3, securing the complex to RNA and helping facilitate transfer of the methyl group to the specific A[119–121].

#### *1.3.1.3 WTAP*

Another core component of the m6A methylation complex is Wilms Tumor Associated Protein (WTAP), identified in 2008 after characterizing the METTL3 homolog in *Arabidopsis*, MTA[115]. MTA was originally identified in a screen resulting in defective plant embryos by Zhong *et al*. [122], and after confirming phenotypes identified MTA as the gene formally called Embryo-Defective1706 (EMB1706). Since human METTL3 was isolated from a complex, the authors also wanted to understand any interacting partners of MTA and performed a yeast two-hybrid screen in tissues highly expressing MTA. This screen resulted in the identification of FIP37, the plant homolog to WTAP. Further characterization of FIP37 proved that it binds to MTA, and resulted in an embryonic lethal phenotype when disrupted. Furthermore, MTA and FIP37 colocalized within nuclear speckles, comparable to the human homologs of these proteins[115].

In yeast, these findings were validated as the yeast homolog to METTL3, Ime4, also binds the yeast homolog of WTAP, mum2[114]. Mum2 was required for m6A formation in yeast, suggesting that the interaction between METTL3 and WTAP is critically important for m6A deposition. In human cells, this suggestion was validated across several cell types and confirmed the function of WTAP to be localization of METTL3 and, by association, METTL14 to nuclear speckles for m6A methylation to occur[97,117,118].

#### *1.3.1.4 VIRMA*

VIRMA (vir-like m6A methyltransferase associated, also known as KIAA1429) is the human homolog of the *Drosophila* protein Virilizer. VIRMA was a top hit in a proteomics study investigating interactors of WTAP, and further proteomic studies also found

similar data[96,97]. Depleting VIRMA resulted in significant loss of m6A, even more than depleting METTL3 or METTL14, indicating its crucial function in m6A deposition[97]. VIRMA is the largest component of the m6A complex (202kDa), and localizes to nuclear speckles like its interacting partner WTAP. VIRMA can recruit the m6A methylation complex to methylate selective regions of the 3' UTR near stop codons in its RNA substrates[123].

#### *1.3.1.5 RBM15/15B*

RBM15 and its paralog RBM15B were identified as high confidence interactors of WTAP via proteomic analyses suggesting that it may be part of the m6A methylation complex[96]. In co-immunoprecipitation experiments, RBM15/15B co-precipitated with METTL3, adding evidence that is a member of the m6A complex. Knockout of WTAP reduced the interaction between RBM15/15B and vice versa, signifying that RMB15/15B interact with METTL3 and the m6A methylation complex in a WTAP dependent manner[124]. Additionally, depletion of RBM15/15B resulted in decreased m6A levels, suggesting RBM15/15B is a key component of the methylation complex.

The most well described function of RBM15/15B is to facilitate the inactivation of *Xist* via m6A methylation by recruiting the METTL3/METTL14/WTAP complex. Knockout of RBM15/15B resulted in decreased *Xist* methylation, preventing *Xist*-mediated transcriptional inhibition[124,125]. Therefore, RBM15/15B is required for proper methylation of *Xist* which promotes transcriptional silencing.

RBM15/15B has also regulates nuclear export of specific RNAs, recognizing RNA transport elements (RTE) and stimulating the export of RTE-containing RNAs by binding directly to mRNA export receptor NXF1[126,127]. Knockout of RBM15/15B lead to

cytoplasmic depletion and nuclear accumulation of mRNAs, suggesting m6A and RBM15/15B have essential functions in mRNA transport.

#### *1.3.1.6 ZC3H13*

Many of the studies that identified m6A components relied upon large scale proteomic experiments to identify interactors of select proteins[128]. As such, there are likely other proteins that interact with m6A complexes and may influence that addition of m6A to RNA. After evaluating endogenous protein complexes across species to look for evolutionarily conserved protein interactions using quantitative mass spectrometry, ZC3H13 was found in complex with WTAP, VIRMA, and Hakai (CBLL1), another protein involved with m6A[129]. In mouse embryonic stem cells (mESCs) Wen et al. identified the c-terminal region to be necessary and sufficient for its interaction with Wtap, Virma, and Hakai[130]. Using tandem mass spectrometry, they found Zc3h13 was critical for m6A methylation, as depletion of Zc3h13 resulted in decreased m6A levels, specifically at 3' UTRs of mRNA. Similar to WTAP, reduction of Zc3h13 decreased nuclear levels of the Zc3h13/Wtap/Virma/Hakai complex, suggesting that Zc3H14 assists in the nuclear localization of the m6A methylation complex[130].

### *1.3.2 m6A demethylases ("erasers")*

#### *1.3.2.1 FTO*

Fat mass and obesity-associated (FTO) was the first protein identified that showed m6A demethylase activity[98]. In HeLa and HEK293 cells, Jia *et al*.[98] showed that knockdown of FTO resulted in a slight increase in m6A and conversely overexpression of FTO slightly reduced levels of m6A. This finding led the authors to believe that FTO

was removing m6A marks, signifying that the m6A modification is dynamic. Given that methylation takes place in the nucleus and m6A writers localize to nuclear speckles, immunofluorescence experiment in HeLa cells showed that FTO partially colocalized with nuclear speckles[98]. However, later studies went on to show that FTO's preferred target is m6Am, rather than m6A[107]. Current understanding of the mechanism proposes that FTO inhibits the addition of new m6A to RNAs[107].

Further exploration into discerning FTO preference for m6A or m6Am comes from Fto knockout mice. Fto knockout mice exhibit alterations in neurotransmission, metabolic rate, and are obese compared to wildtype mice[131,132]. The link to obesity was intriguing as this was the first indication that m6A could be connected to disease. However, MeRIP-seq analysis of Fto knockout mice only showed changes in a subset of m6A sites throughout the transcriptome when a more global increase was expected[132]. Interestingly, the m6A peaks that showed the highest degree of change were in the 5' UTR, and more detailed analysis proved that these were m6Am sites rather than m6A[107].

#### *1.3.2.2 ALKBH5*

ALKBH5 was discovered after a biochemical screen looking for enzymes that exhibit demethylase activity[95]. Validation proved that ALKBH5 was able to demethylate a ssRNA probe *in vitro*, and knockout of ALKBH5 resulted in mildly increased m6A levels. The most notable change in m6A levels came from overexpression of ALKBH5, which significantly reduced m6A abundance. Even though FTO was believed to be a demethylase at first, its partial colocalization with nuclear speckles was suspicious of activity in removing m6A marks. Conversely, ALKBH5 fully colocalizes with nuclear

speckles in an RNA dependent manner, and loss of ALKBH5 reduced levels of nuclear speckle markers, providing more confidence that it is a true demethylase[95].

Despite activity specific for m6A, the function of ALBKH5 as a demethylase are not fully clear. In one study, ALKBH5 overexpression reduced *NANOG* mRNA m6A levels in breast cancer cells, but subsequent studies determined to be a result of decreased methyltransferase activity[133,134]. Unlike FTO-knockout mice, ALKBH5 knockout mice show no real phenotypic difference compared to controls, except in spermatogenesis[95]. These data suggest that ALKBH5 may not be a global demethylase, but rather may prevent methylation from occurring in the first place.

Insight into how ALKBH5 may function came when Zhang et al. identified an RNA target that seemed to be regulated by ALKBH5[135]. *FOXM1* is demethylated in the context of an RNA-RNA interaction with its antisense transcript *AS-FOXM1*[135]. When this RNA-RNA interaction forms, it may create a structure that recruits ALKBH5 and allows for demethylation. Similarly to how m6A writers can prefer certain structures[108], this is potentially the case for ALKBH5, opening up potential avenues for further investigation of RNA-RNA interactions and their respective influence on methylation.

### *1.3.3 m6A modification binding proteins "readers"*

#### *1.3.3.1 YTHDF1-3*

Whichever m6A RNA binding protein binds to it methylated RNAs determine the effect of m6A on that RNA transcript. These "readers" thereby influence the fate of that RNA, taking the transcript to be translated, spliced, degraded, or exported. Of the m6A readers, YTH domain-containing family proteins (YTHDF) are the most prominent. Initially discovered from m6A immunoprecipitation experiments, proteins with a YT521-B

homologous domain were among the strongest interactors with m6A[103]. This study identified YTHDF2 and YTHDF3 as m6A binding proteins, and future *in vitro* studies validated that the YTH domain binds RNA in an m6A dependent manner[136,137]. Along with another family member YTHDF1, these three proteins are highly similar and share nearly the same amino acid sequence. Other than the YTH domain, the rest of the YTHDF proteins are mostly low complexity and enriched with Q, N, and P residues. Being largely disordered, YTHDF proteins phase separate especially when bound to m6A RNA[138]. Like other RNA binding proteins that phase separate, YTHDF proteins localize to stress granules and P-bodies with bound m6A RNA[138]. YTHDF proteins are located in the cytoplasm, where they bind to methylated RNAs found in the DRACH motif[92,139,140].

There is some controversy surrounding the function of YTHDF proteins, and how they affect RNA processing. Initial experiments stated that each YTHDF protein had a different effect on the RNA: YTHDF1 enhanced translation of methylated RNAs, YTHDF2 enhanced their degradation, and YTHDF3 was suggested to be a stabilizer for both and help with their function[92,139,140]. However, other studies suggested that all three are involved with RNA stability, and recognition of m6A RNA would promote degradation of that trasncript[141,142]. Adding to the discrepancies for the YTHDF proteins are conflicting studies describing the selectivity of RNA binding. One study suggest that RNAs are bound by one of the reader proteins and not the others[140], while another suggests that since the proteins are very similar they share many of the same RNA targets[124]. Regardless, more studies are needed to determine the functionality of these proteins and understand their contributions to the m6A pathway.

#### *1.3.3.2 YTHDC1-2*

Other than the YTHDFs, there are two other YTH containing proteins, YTHDC1 and YTHDC2. YTHDC1 localizes to the nucleus, where it is believed to function in splicing, export, and silencing the lncRNA *Xist*[124,143–145]. Specifically, it localizes to nuclear speckles, areas of transcription and splicing, and also where the writers and erasers are localized[144]. YTHDC1's association with nuclear speckles may indicate that YTHDC1 binds RNAs shortly after they are transcribed to exert its splicing capabilities. In fact, YTHDC1 binding of *SRSF3* leads to enhanced exon inclusion, likely as a result of displacing SRSF10[144]. Localization in the nucleus also allows binding to lncRNAs that often act as scaffolds for nuclear bodies. As such, YTHDC1 was shown using CLIP experiments to preferentially bind lncRNAs like *XIST*, *NEAT1*, and *MALAT1*[124]. Specifically with *XIST*, YTHDC1 binds to the many m6A sites, and is required for XIST-mediated silencing in an m6A dependent manner[124]. Additional studies are needed to explore the function of YTHDC1 in terms of other lncRNAs or mRNAs to fully understand its role in m6A recognition.

YTHDC2 is an m6A binding protein that has a few unique characteristics separating it from YTHDC1 and the YTHDFs. Firstly, YTHDC2 is only expressed in the testes, whereas the others are ubiquitously expressed throughout the body[146–149]. Here, it is critical for spermatogenesis and reproduction as mice lacking YTHDC2 are infertile, and both male and female have significantly smaller gonads[146–149]. YTHDC2 also exhibits a decreased propensity to bind m6A compared to the other YTH proteins[146,150]. The weaker binding to m6A directly may be explained by the slight differences that YTHDC2 has in its m6A binding region, supported by CLIP data where YTHDC2 has

less overlap with m6A sites than other YTH proteins[124,151]. Thirdly, YTHDC2 has far more domains than the other YTH proteins, one of which being an RNA helicase domain. The RNA helicase domain could be useful for promoting translation by unwinding RNA so that it is more accessible for translation machinery. Conflicting data makes it difficult to understand the exact function of YTHDC2 because some data indicates increased translation with YTHDC2 expression[148] but other data demonstrates mRNA degradation through recruitment of exoribonuclease Xrn1 and higher expression of m6A containing transcripts with YTHDC2 depletion[146,152]. Like the other YTH proteins, more data are needed to determine YTHDC2 function, especially given its different protein domains and behaviors.

#### *1.3.3.3 IGF2BP1-3*

A YTH domain is not the only domain capable of binding m6A, as proteins such as the insulin growth factor 2 mRNA binding proteins (IGFBP) lack a YTH domain but still bind m6A[153]. IGF2BP proteins were identified as m6A binding proteins via immunoprecipitation experiments using a m6A or non-m6A probe followed by mass spectrometry[153]. IGF2BP proteins differ from YTH containing proteins in that they enhance stability of bound transcripts[153]. IGF2BP's have a highly similar consensus sequence (UGGAC) to DRACH which suggests they may bind m6A. IGF2BP proteins bind YTHDF proteins in pulldown studies and m6A binding may rely on the m6A switch mechanism[154,155].

#### *1.3.3.4 HNRNP proteins*

Another group of proteins that bind m6A RNAs are HNRNPs. They bind methylated RNAs not in a direct way, but as a result of the structural changes brought on by m6A addition. RNA often forms secondary structures like helices, stem loops, or hairpins, and m6A sites within these regions can alter the stability of these structures. Changes in stability result from differences in base pair strength because m6A-U base pairs are weaker than A-U. An m6A mark can reduce the melting temperature by 5°C or more, forcing the RNA into a single-stranded conformation instead of a structured, double-stranded one[156]. In fact, transcriptome wide characterization of RNA structure determined that m6A often falls in unstructured areas, perhaps destabilized by the m6A modification[157]. This phenomenon was termed a “m6A structural switch” or “m6A switch”[108]. HNRNPC first demonstrated this phenomenon after studies determined that an m6A site within a hairpin changed the confirmation so that HNRNPC could now bind its U-rich consensus motif with a newly formed single stranded RNA[108]. HNRNPC is involved with pre-mRNA processing, specifically splicing, stability, expression, of which can be influenced by the presence or absence of m6A[108]. Likewise, m6A can enhance binding of proteins with binding sites nearby to an occupied m6A. This is the case for two other HNRNP proteins, HNRNPG and HNRNPA2B1 which exhibit decreased binding after loss of m6A near their binding sites[158,159].

Direct and indirect readers of m6A have important functions in many stages of RNA metabolism that are influenced by the presence of m6A and RNA structure. Since the two are linked, it is sometimes difficult to separate if a protein can recognize m6A on

its own, or if the changes in structure allow the protein to bind. More work is needed in this area to characterize m6A readers and identify previously unknown readers.

### *1.3.4 Consequences of m6A modification*

#### *1.3.4.1 RNA stability*

Perhaps the most well-known effect of m6A on RNA is the impact it has on RNA stability. Around the time when m6A was discovered, half-life measurements of mRNAs using radioisotope labeling with or without m6A determined that mRNAs containing m6A had a shorter half life[160]. Despite this finding, the destabilizing effects of m6A were not investigated for almost 40 years. The next finding that m6A influences mRNA stability came from studies that lowered levels of m6A writer proteins METTL3 or WTAP. When these proteins were knocked down, an increase in mRNA half-life was observed in both human and mouse[97,117,161]. Of the m6A readers, YTHDF2 was identified to be the primary reader involved in RNA degradation[92]. After identifying binding partners of YTHDF2, Wang *et. al*[162] then asked whether depleting YTHDF2 could change mRNA levels after hypothesizing that YTHDF2 is involved in stability because YTHDF2 primarily binds to RNA substrates in the 3' UTR. Indeed, knockdown of YTHDF2 increased half-life by 30% and had a more pronounced effect based on the amount of m6A sites the transcript had[92]. This provided the first evidence that m6A induced instability because of direct protein recognition. Other evidence has suggested that all three YTHDF proteins function to destabilize RNA, and that it is not unique to YTHDF2[140,163,164]. However, YTHDF2 involvement in RNA stability has been more characterized than the other two proteins, addressed below.

Recent explorations into how YTHDF2 degrades methylated transcripts has uncovered a few pathways that are involved with RNA decay. Early studies showed that YTHDF2 localized RNAs to processing bodies (P-bodies), which are liquid like cytoplasmic granules that hold proteins responsible for RNA decay[92,138,165]. YTHDF2 and other YTHDF family proteins also localize to stress granules, and recent work suggests that YTHDF family proteins may promote stress granule formation[140,162,166,167]. While stress granules don't necessarily degrade RNAs in them, m6A modified RNAs are enriched in stress granules. Dynamic regulation of these RNAs by m6A suggests that changes in stability could alter normal cellular function under times of stress[168].

Direct RNA destabilization via YTHDF2 occurs through the recruitment of the CCR4-NOT deadenylation complex. Through an interaction on the N-terminus, YTHDF2 directly binds to the SH domain of CNOT1, and this interaction is essential for deadenylation of RNAs[169]. This was the first evidence of an associated degradation pathway tied to YTHDF2 function in RNA degradation, and important also because the CCR4-NOT complex acts before localizing deadenylated RNAs to P-bodies[170], suggesting an alternative mechanism for YTHDF2 that may precede its involvement in P-bodies. An alternative pathway for YTHDF2-dependent degradation is through the RNase P/MRP pathway[171]. Here, YTHDF2 binds to an adaptor protein HRSP12 to connect bound m6A RNAs to the endoribonuclease P/MRP where they are cleaved and degraded. HRSP12 also facilitates YTHDF2 binding to RNAs and vice versa, but more understanding of this joint facilitation of binding is needed to conclude mechanism and future implications. Interestingly, RNA substrates not degraded by HRSP12-P/MRP endoribonucleolytic cleavage heavily depend upon upstream frameshift 1 (UPF1), a key

regulator of nonsense-mediated decay, to destabilize mRNA[172]. Destabilization of m6A RNA relies on a physical interaction with UPF1 at the N-terminus of YTHDF2, proper ATPase/helicase activities of UPF1, and interaction of UPF1 with a decapping promoting factor proline-rich nuclear receptor coactivator 2 (PNRC2)[172]. These connections to independent RNA degradation pathways secure YTHDF2 as a modifier of RNA stability and degradation through m6A recognition.

#### *1.3.4.2 mRNA splicing*

The addition of m6A to RNA co-transcriptionally is one mechanism in which m6A can control the splicing of RNA transcripts. Perhaps the most well characterized m6A regulated splicing event is that of the master sex determining factor Sex-lethal in *Drosophila*. In flies, Female-lethal(2)d (Fl(2)d) and Virilizer are required for sex dependent regulation of *Sex-lethal* (*Sxl*). *Sxl* is regulated by alternative splicing from an intronic m6A site that is recognized by the fly homolog of YTHDF proteins, YT521-B[173–175]. Knockout of METTL3 homolog *Ime4* prevents proper female-specific splicing of *Sxl* resulting in flies that were predominantly male. Proper splicing of *Sxl* also represses *male-specific lethal* (*msl-2*) to prevent dosage compensation in females. While somewhat well characterized in flies, there are much less data on how m6A influences splicing in human models. For example, instances of METTL3-mediated methylation at exonic or intronic splice sites are rare. In METTL3-knockout embryonic stem cells, ~0.15-0.5% of alternatively spliced exons had altered splicing compared to control cells[116,176]. Another study using METTL3-knockout cells also found very few transcripts with splicing changes, and of the transcripts with splicing changes only a third even contained m6A[176]. Although METTL3 knockout had little effect on a global scale of

alternative splicing, there are examples where changes in methylation alter splicing and these situations are often under cellular stress. For instance, endoplasmic reticulum stress (ER) and innate immunity sensing following *Flaviviridae* decreased the amount of m6A in *CIRBP* mRNA[177]. The infection decreased intron retention of a m6A-containing long intron within a specific isoform of CIRBP. The loss of the long intron resulted in less *CIRBP* mRNA and protein, which subsequently led to a decrease in the production of infectious viruses. This mechanism suggests m6A can alter splicing, potentially in a manner of self-preservation against infecting viruses. Another example of the importance of m6A in splicing is observed with TARBP2. This RNA binding protein regulates RNA stability within the nucleus by binding and retaining introns of bound pre-mRNA substrates that are later sent to be degraded by the nuclear exosome. TARBP2 recruits the methylation component WTAP to methylate introns, and the resulting methylation repels certain splicing factors, like SRSF1, from efficiently splicing out the intron[178]. Another m6A methyltransferase, METTL16, relies upon m6A to rapidly induce splicing of an intron from *MAT2A*, which encodes a SAM synthase. Regulated splicing of *MAT2A* maintains consistent levels of SAM within the cell, and under SAM-deficient conditions, METTL16 binds to its substrate vertebrate conserved hairpin (hp1) located in the 3' UTR of MAT2A to induce splicing. As a result, MAT2A expression is upregulated and more SAM is created, until the point where there are sufficient levels of SAM which allows METTL16 to methylate an m6A site within hp1, inducing intron retention and subsequent degradation[179].

Despite few RNAs being directly influenced by METTL3 depletion, m6A is more common on dynamically regulated RNAs than housekeeping genes[97], suggesting that it

may be more important than we are currently able to determine. A significant result of splicing that may be methylation dependent is the proteins that are recruited or repelled. The presence of m6A can recruit or repel specific proteins, of which many are splicing factors and regulators of RNA stability, one being TDP43[180]. Though m6A may not be directly altering the splicing of an intron, secondary recruitment of other proteins is a possibility that cannot be overlooked.

#### *1.3.4.3 mRNA export*

As the next step in RNA processing, m6A modifications prevent or enhance export of RNAs from the nucleus to be translated in the cytoplasm. After infection, RNA helicase DDX46 recruits ALKBH5 to demethylate m6A modified transcripts important for producing type I interferons. Once these RNAs are demethylated, they are entrapped within the nucleus and the antiviral innate immunity response is suppressed[181]. Conversely, in forms of blood cancer the lncRNA *MALAT1*, known to act as a scaffold for nuclear speckles, promotes the export of chimeric mRNA and fusion proteins in an m6A dependent manner by recruiting the METTL3-METTL14 complex via METTL14. By recruiting this complex, *MALAT1* forces interactions between oncogenic fusion proteins and the METTL3-METTL14 complex, resulting in suppressed hematopoietic cell differentiation[182].

An additional consequence of m6A modification is facilitating nuclear export of mature mRNAs via specific interactions. Methylated transcripts bound by YTHDC1 are exported from the nucleus after delivery to nuclear mRNA export receptor NXF1 and interactions with splicing factor and export adaptor SRSF3[145]. FMRP, another m6A reader protein that enhances stability of its targets[183], exports RNAs through the CRM1-

dependent export pathway during neural differentiation[184,185]. These interactions help facilitate the critical export of m6A-containing mRNAs into the cytoplasm so that they are translated into protein.

#### *1.3.4.4 mRNA Translation*

The impact of m6A on translation ultimately depends on which reader protein binds to the RNA, interactions between methylation components and factors involved with the translation mechanisms, and where on the RNA the m6A mark is. The rate of transcription determines translation efficiency, and some data suggests that slower transcription leads to more m6A per RNA[186], although how this affects the rate of translation is still being investigated. Likely, the location of m6A will be the most important factor in determining which readers recognize it or how quickly it is translated. METTL3 promotes translation by recognizing 5' UTR m6A and 3' UTR m6A. For eukaryotic cells, protein synthesis usually starts with eIF4F binding to the m7G cap found at the 5' end of mRNAs. The eIF4F complex is composed of a 5' cap binding protein (eIF4E), an RNA helicase (eIF4A), and a scaffold protein (eIF4G) that binds eIF4E and eIF4A together[187]. However, with methylation at the 5' end, translation can occur independent of eIF4F and instead by binding to initiation factors such as eIF3 in an m6A-regulated manner[188]. A theory for how METTL3 regulates translation is by binding to mRNA after methylating it where it is then exported to the cytoplasm with the RNA and interacting with eIF3. This interaction facilitates additional interactions with mRNA cap associated proteins, forming an mRNA loop. This loop would allow METTL3 to interact with m6A at both the 5' and 3' end, promoting eIF3 function at the 5' end. As such, this proposed looping mechanism can only promote translation when METTL3

binds near the stop codon in 3' UTR, and this approach has also been proposed to explain translation promotion by ribosome cycling[189].

The first reader protein identified to impact translation was YTHDF1. Knockdown of YTHDF1 in HeLa cells reduced translation efficiency of target RNAs determined by ribosome profiling[139]. Additionally, YTHDF1 knockdown increased the amount of m6A methylated RNAs in the non-translatable pool and decreased the amount in the translatable pool. The translation efficiency of YTHDF1 targets were also decreased by METTL3 knockdown implicating the m6A recognition as a key component for successful translation of target RNAs. YTHDF1 binds eIF3 to recruit the translation initiation complex to target RNAs and enhance translation, although the exact mechanism for how this enhances translation is unclear[139]. Somewhat controversial, m6A's influence on translation could have significant global effects that will need to be validated through continued studies.

#### *1.3.4.5 m6A in neurodegenerative disease*

Alterations of m6A homeostasis have been connected to several different diseases now that recent methods allow investigation into changes in methylation levels and sites. m6A has been extensively studied in various forms of cancer and changes in m6A writers, erasers, or readers being implicated in cancer progression[190]. However, the importance of m6A in neurodegeneration is still relatively unknown despite the brain having the highest levels of methylation and m6A being critical for nervous system development[93,132,174,191]. Recent studies have investigated the function of m6A in some forms of neurodegeneration such as Alzheimer's Disease (AD) and Parkinson's Disease (PD). In rat brain and PC12 cells treated with 6-hydroxydopamine (6-OHDA), a drug to

stimulate calcium influx and oxidative stress by inducing NMDA receptor expression, there was death of dopaminergic neurons that eventually resulted in PD like phenotypes[192]. Cells under oxidative stress exhibit altered m6A profiles, and more m6A can be found on the 5' end of RNAs[168]. Additionally, analysis of two PD datasets suggests that m6A reader HNRNPC is downregulated in PD, and may contribute to disease pathogenesis[193]. Although promising, more data are needed to conclude mechanistic involvement of m6A in PD that would create opportunities for therapeutic targeting.

Studies connecting AD to changes in methylation have also yielded some results that may implicate m6A in disease. In one instance, METTL3 expression in the cortex and hippocampus of AD mice models was upregulated, while FTO was decreased, resulting in higher levels of m6A present than compared to controls[194]. However, a more recent study in human AD brains found that there was decreased m6A, and decreased expression of METTL3 compared to controls. The authors also found reduced m6A after ablating METTL3 in the hippocampus led to memory deficits, dramatic synapse loss, and neuronal death. Accompanying these effects were cellular states mimicking AD, like oxidative stress and aberrant cell cycle events. Notably, overexpression of METTL3 rescued these effects[195]. Genome wide association studies (GWAS) identified two variants of FTO that may increase AD risk, and also found that FTO interacts with apolipoprotein E (APOE) to increase the risk of dementia[196]. Other mechanistic studies investigating how FTO may be involved with AD determined that FTO activates the TSC1-mTOR-Tau signaling pathway by reducing m6A levels. In concordance with this, FTO expression is increased in brains of transgenic AD mice and knocking down FTO

in these models is sufficient to improve cognitive ability[192,197]. Commonly associated with AD is aggregates or oligomers of microtubule-associated protein tau. Using a Cry2-based optogenetic model to induce tau oligomerization, Jiang *et al.*[198] identified m6A reader HNRNPA2B1 as a primary target of oligomerized tau (oTau) via proteomic analysis. The authors went on to confirm the interaction between oTau and HNRNPA2B1 in neurons, mouse models, and human AD brain samples[198]. Knockout of HNRNPA2B1 prevented oTau association with m6A and reduced oTau-mediated neurodegeneration. In addition, AD brains display 5-fold increased levels of the m6A-oTau-HNRNPA2B1 complex. Jiang *et al.*[198] theorized that HNRNPA2B1 acts as a linker between oTau and m6A methylated transcripts, which may drive Tau pathology in AD.

These findings suggest m6A may contribute to disease mechanisms, but interpretation is complicated by the significant differences amongst the findings. More extensive characterization is needed to illicit the mechanistic alterations occurring in AD and PD and to what degree they are a result of m6A dysregulation.

### 1.4 Summary and dissertation goals

TDP43 is an essential protein for maintaining RNA homeostasis with active roles in RNA splicing, transport, miRNA biogenesis, and stress granule formation, culminating in an overall theme of regulating RNA stability. In ALS, nuclear TDP43 is mislocalized to the cytoplasm, and this shift is a hallmark of pathology. Despite many advances over time in understanding the molecular events that lead to ALS, TDP43 function and the mechanism for its mislocalization remain unknown. It is not clear whether ALS results

from a loss of function of TDP43, or a gain of function, or if there is another disease mechanism altogether.

Previous data from our lab described widespread RNA instability with TDP43 overexpression and in ALS, specifically involving RNAs encoding oxidative phosphorylation enzymes and ribosomal proteins. Furthermore, evidence suggests that these transcripts may be regulated by m6A modifications. Koranda *et. al*[199] found that knock out of m6A methyltransferase, METTL14, similarly resulted in upregulated RNA levels of mitochondrial and protein encoding ribosomal RNAs[199]. These findings suggest that there may be a relationship between the RNAs destabilized by TDP43 and m6A. The goal of this thesis is to explore this very question, aiming to establish TDP43 as an RNA binding protein that recognizes m6A-containing RNAs and understand to what degree methylation has on the functions of TDP43. Chapter 2 identifies the methylation status of TDP43 RNA substrates and how alterations to m6A can alter the propensity of TDP43 to bind RNA. In addition, I assay the methylation levels in ALS spinal cord, as well as identify m6A factors that can rescue TDP43-dependent toxicity in disease models. Chapter 3 provides context to some of the findings, and a discussion of unanswered questions for the ALS and m6A fields that should lead research in the upcoming years. Lastly, Appendix A outlines the development of a CRISPR based tool to study knockdown of multiple genes at once, or enable creation of CRISPR edited cell models.

## References


1. Charcot J., Joffory A. Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arcives Physiol Neurol Pathol*. 1869;2:744-754.
2. Talbott EO, Malek AM, Lacomis D. The epidemiology of amyotrophic lateral sclerosis. *Handb Clin Neurol*. 2016;138:225-238. doi:10.1016/B978-0-12-802973-2.00013-6
3. Chia R, Chiò A, Traynor BJ. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *Lancet Neurol*. 2018;17(1):94-102. doi:10.1016/S1474-4422(17)30401-5
4. Ferrari R, Kapogiannis D, Huey ED, Momeni P. FTD and ALS: a tale of two diseases. *Curr Alzheimer Res*. 2011;8(3):273-294. http://www.ncbi.nlm.nih.gov/pubmed/21222600.
5. Neumann M. Molecular neuropathology of TDP-43 proteinopathies. *Int J Mol Sci*. 2009;10(1):232-246. doi:10.3390/ijms10010232
6. Ou SH, Wu F, Harrich D, García-Martínez LF, Gaynor RB. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol*. 1995;69(6):3584-3596. http://www.ncbi.nlm.nih.gov/pubmed/7745706. Accessed December 18, 2018.
7. Shiina Y, Arima K, Tabunoki H, Satoh JI. TDP-43 dimerizes in human cells in culture. *Cell Mol Neurobiol*. 2010;30(4):641-652. doi:10.1007/S10571-009-9489-9/FIGURES/8
8. Zhang YJ, Caulfield T, Xu YF, et al. The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum Mol Genet*. 2013;22(15):3112-3122. doi:10.1093/HMG/DDT166
9. Ayala YM, Zago P, D'Ambrogio A, et al. Structural determinants of the cellular localization and shuttling of TDP-43. *J Cell Sci*. 2008;121(Pt 22):3778-3785. doi:10.1242/JCS.038950
10. Winton MJ, Igaz LM, Wong MM, Kwong LK, Trojanowski JQ, Lee VMY. Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *J Biol Chem*. 2008;283(19):13302-13309. doi:10.1074/JBC.M800342200
11. Kuo P-H, Doudeva LG, Wang Y-T, Shen C-KJ, Yuan HS. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res*. 2009;37(6):1799-1808. doi:10.1093/nar/gkp013
12. Ayala YM, Pantano S, D'Ambrogio A, et al. Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J Mol Biol*. 2005;348(3):575-588. doi:10.1016/J.JMB.2005.02.038
13. Buratti E, Baralle FE. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator ofCFTR Exon 9 *. *J Biol Chem*. 2001;276(39):36337-36343. doi:10.1074/JBC.M104236200
14. Kuo PH, Chiang CH, Wang YT, Doudeva LG, Yuan HS. The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res*. 2014;42(7):4712. doi:10.1093/NAR/GKT1407

15. Lukavsky PJ, Daujotyte D, Tollervey JR, et al. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol*. 2013;20(12):1443-1449. doi:10.1038/NSMB.2698
16. D'Ambrogio A, Buratti E, Stuani C, et al. Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Res*. 2009;37(12):4116-4126. doi:10.1093/NAR/GKP342
17. Buratti E, Brindisi A, Giombi M, Tisminetzky S, Ayala YM, Baralle FE. TDP-43 Binds Heterogeneous Nuclear Ribonucleoprotein A/B through Its C-terminal Tail: AN IMPORTANT REGION FOR THE INHIBITION OF CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR EXON 9 SPLICING. *J Biol Chem*. 2005;280(45):37572-37584. doi:10.1074/JBC.M505557200
18. Deshaies J-E, Shkreta L, Moszczynski AJ, et al. TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis. *Brain*. 2018;141(5):1320-1333. doi:10.1093/brain/awy062
19. Pesiridis GS, Lee VM-Y, Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009;18(R2):R156-62. doi:10.1093/hmg/ddp303
20. Polymenidou M, Lagier-Tourenne C, Hutt KR, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14(4):459-468. doi:10.1038/nn.2779
21. Flores BN, Li X, Malik AM, Martinez J, Beg AA, Barmada SJ. An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep*. 2019;27(4):1133-1150.e8. doi:10.1016/j.celrep.2019.03.093
22. Elden AC, Kim HJ, Hart MP, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nat 2010 4667310*. 2010;466(7310):1069-1075. doi:10.1038/nature09320
23. Voigt A, Herholz D, Fiesel FC, et al. TDP-43-Mediated Neuron Loss In Vivo Requires RNA-Binding Activity. *PLoS One*. 2010;5(8):e12247. doi:10.1371/JOURNAL.PONE.0012247
24. Ihara R, Matsukawa K, Nagata Y, et al. RNA binding mediates neurotoxicity in the transgenic Drosophila model of TDP-43 proteinopathy. *Hum Mol Genet*. 2013;22(22):4474-4484. doi:10.1093/HMG/DDT296
25. Coyne AN, Yamada SB, Siddegowda BB, et al. Fragile X protein mitigates TDP-43 toxicity by remodeling RNA granules and restoring translation. *Hum Mol Genet*. 2015;24(24):6886. doi:10.1093/HMG/DDV389
26. Lu Y, Ferris J, Gao FB. Frontotemporal dementia and amyotrophic lateral sclerosis-associated disease protein TDP-43 promotes dendritic branching. *Mol Brain*. 2009;2(1):1-10. doi:10.1186/1756-6606-2-30/FIGURES/6
27. Johnson BS, McCaffery JM, Lindquist S, Gitler AD. A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proc Natl Acad Sci U S A*. 2008;105(17):6439-6444. doi:10.1073/PNAS.0802082105/SUPPL_FILE/0802082105SI.PDF
28. Li Y, Ray P, Rao EJ, et al. A Drosophila model for TDP-43 proteinopathy. *Proc Natl Acad Sci U S A*. 2010;107(7):3169-3174. doi:10.1073/PNAS.0913602107/SUPPL_FILE/PNAS.200913602SI.PDF

29. Estes PS, Boehringer A, Zwick R, Tang JE, Grigsby B, Zarnescu DC. Wild-type and A315T mutant TDP-43 exert differential neurotoxicity in a Drosophila model of ALS. *Hum Mol Genet*. 2011;20(12):2308-2321. doi:10.1093/HMG/DDR124
30. Estes PS, Daniel SG, Mccallum AP, et al. Motor neurons and glia exhibit specific individualized responses to TDP-43 expression in a Drosophila model of amyotrophic lateral sclerosis. *Dis Model Mech*. 2013;6(3):721. doi:10.1242/DMM.010710
31. Liachko NF, Guthrie CR, Kraemer BC. Phosphorylation Promotes Neurotoxicity in a Caenorhabditis elegans Model of TDP-43 Proteinopathy. *J Neurosci*. 2010;30(48):16208-16219. doi:10.1523/JNEUROSCI.2911-10.2010
32. Barmada SJ, Skibinski G, Korb E, Rao EJ, Wu JY, Finkbeiner S. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci*. 2010;30(2):639-649. doi:10.1523/JNEUROSCI.4988-09.2010
33. Walker AK, Spiller KJ, Ge G, et al. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. *Acta Neuropathol*. 2015;130(5):643. doi:10.1007/S00401-015-1460-X
34. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet 2008 4012*. 2008;40(12):1413-1415. doi:10.1038/ng.259
35. Johnson MB, Kawasawa YI, Mason CE, et al. Functional and Evolutionary Insights Into Human Brain Development Through Global Transcriptome Analysis. *Neuron*. 2009;62(4):494. doi:10.1016/J.NEURON.2009.03.027
36. Yeo G, Holste D, Kreiman G, Burge CB. Variation in alternative splicing across human tissues. *Genome Biol*. 2004;5(10):R74. doi:10.1186/GB-2004-5-10-R74
37. Tollervey JR, Curk T, Rogelj B, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci*. 2011;14(4):452-458. doi:10.1038/nn.2778
38. Sephton CF, Cenik C, Kucukural A, et al. Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J Biol Chem*. 2011;286(2):1204-1215. doi:10.1074/jbc.M110.190884
39. Ling JP, Chhabra R, Merran JD, et al. PTBP1 and PTBP2 repress nonconserved cryptic exons. *Cell Rep*. 2016;17(1):104. doi:10.1016/J.CELREP.2016.08.071
40. Tan Q, Yalamanchili HK, Park J, et al. Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum Mol Genet*. 2016;25(23):5083-5093. doi:10.1093/hmg/ddw337
41. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A*. 2003;100(1):189. doi:10.1073/PNAS.0136770100
42. Jeong YH, Ling JP, Lin SZ, et al. Tdp-43 cryptic exons are highly variable between cell types. *Mol Neurodegener*. 2017;12(1). doi:10.1186/S13024-016-0144-X
43. Ling JP, Pletnikova O, Troncoso JC, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 2015;349(6248):650-655. doi:10.1126/science.aab0983
44. Li Q, Zheng S, Han A, et al. The splicing regulator PTBP2 controls a program of

embryonic splicing required for neuronal maturation. *Elife*. 2014;2014(3):1201. doi:10.7554/ELIFE.01201.001
45. Suckale J, Wendling O, Masjkur J, et al. PTBP1 Is Required for Embryonic Development before Gastrulation. *PLoS One*. 2011;6(2). doi:10.1371/JOURNAL.PONE.0016992
46. Kraemer BC, Schuck T, Wheeler JM, et al. Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis. *Acta Neuropathol*. 2010;119(4):409. doi:10.1007/S00401-010-0659-0
47. Wu LS, Cheng WC, Shen CKJ. Targeted Depletion of TDP-43 Expression in the Spinal Cord Motor Neurons Leads to the Development of Amyotrophic Lateral Sclerosis-like Phenotypes in Mice. *J Biol Chem*. 2012;287(33):27335. doi:10.1074/JBC.M112.359000
48. Iguchi Y, Katsuno M, Niwa JI, et al. Loss of TDP-43 causes age-dependent progressive motor neuron degeneration. *Brain*. 2013;136(Pt 5):1371-1382. doi:10.1093/BRAIN/AWT029
49. Bhardwaj A, Myers MP, Buratti E, Baralle FE. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res*. 2013;41(9):5062-5074. doi:10.1093/nar/gkt189
50. Ayala YM, De Conti L, Avendaño-Vázquez SE, et al. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J*. 2011;30(2):277-288. doi:10.1038/emboj.2010.310
51. Xu Y-F, Gendron TF, Zhang Y-J, et al. Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J Neurosci*. 2010;30(32):10851-10859. doi:10.1523/JNEUROSCI.1630-10.2010
52. Masters CL, Kril JJ, Halliday GM, et al. Overview and recent advances in neuropathology. Part 2: Neurodegeneration. *Pathology*. 2011;43(2):93-102. doi:10.1097/PAT.0B013E3283426EEE
53. Mutihac R, Alegre-Abarrategui J, Gordon D, et al. TARDBP pathogenic mutations increase cytoplasmic translocation of TDP-43 and cause reduction of endoplasmic reticulum Ca2 + signaling in motor neurons. *Neurobiol Dis*. 2015;75:64-77. doi:10.1016/J.NBD.2014.12.010
54. Mitsuzawa S, Akiyama T, Nishiyama A, et al. TARDBP p.G376D mutation, found in rapid progressive familial ALS, induces mislocalization of TDP-43. *eNeurologicalSci*. 2018;11:20-22. doi:10.1016/J.ENSCI.2018.04.001
55. Melamed Z, López-Erauskin J, Baughn MW, et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*. 2019;22(2):180-190. doi:10.1038/s41593-018-0293-z
56. Klim JR, Williams LA, Limone F, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci*. January 2019:1. doi:10.1038/s41593-018-0300-4
57. Brown AL, Wilkins OG, Keuss MJ, et al. TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. *Nat 2022 6037899*. 2022;603(7899):131-137. doi:10.1038/s41586-022-04436-3
58. Ma XR, Prudencio M, Koike Y, et al. TDP-43 represses cryptic exon inclusion in

the FTD–ALS gene UNC13A. *Nat 2022 6037899*. 2022;603(7899):124-130. doi:10.1038/s41586-022-04424-7
59. Tank EM, Figueroa-Romero C, Hinder LM, et al. Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat Commun*. 2018;9(1):2845. doi:10.1038/s41467-018-05049-z
60. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. *Neuron*. 2011;72(2):245. doi:10.1016/J.NEURON.2011.09.011
61. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011;72(2):257. doi:10.1016/J.NEURON.2011.09.010
62. Kedersha NL, Gupta M, Li W, Miller I, Anderson P. RNA-Binding Proteins Tia-1 and Tiar Link the Phosphorylation of Eif-2α to the Assembly of Mammalian Stress Granules. *J Cell Biol*. 1999;147(7):1431. doi:10.1083/JCB.147.7.1431
63. Harding HP, Novoa I, Zhang Y, et al. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell*. 2000;6(5):1099-1108. doi:10.1016/S1097-2765(00)00108-8
64. Mazroui R, Sukarieh R, Bordeleau ME, et al. Inhibition of Ribosome Recruitment Induces Stress Granule Formation Independently of Eukaryotic Initiation Factor 2α Phosphorylation. *Mol Biol Cell*. 2006;17(10):4212. doi:10.1091/MBC.E06-04-0318
65. Colombrita C, Zennaro E, Fallini C, et al. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem*. 2009;111(4):1051-1061. doi:10.1111/J.1471-4159.2009.06383.X
66. Dewey CM, Cenik B, Sephton CF, et al. TDP-43 Is Directed to Stress Granules by Sorbitol, a Novel Physiological Osmotic and Oxidative Stressor. *Mol Cell Biol*. 2011;31(5):1098. doi:10.1128/MCB.01279-10
67. Liu-Yesucevitz L, Bilgutay A, Zhang YJ, et al. Tar DNA Binding Protein-43 (TDP-43) Associates with Stress Granules: Analysis of Cultured Cells and Pathological Brain Tissue. *PLoS One*. 2010;5(10):13250. doi:10.1371/JOURNAL.PONE.0013250
68. McDonald KK, Aulas A, Destroismaisons L, et al. TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum Mol Genet*. 2011;20(7):1400-1410. doi:10.1093/HMG/DDR021
69. Aulas A, Velde C Vande. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? *Front Cell Neurosci*. 2015;9(OCTOBER). doi:10.3389/FNCEL.2015.00423
70. Khalfallah Y, Kuta R, Grasmuck C, Prat A, Durham HD, Vande Velde C. TDP-43 regulation of stress granule dynamics in neurodegenerative disease-relevant cell types. *Sci Rep*. 2018;8(1). doi:10.1038/S41598-018-25767-0
71. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*. 2004;10(12):1957. doi:10.1261/RNA.7135204
72. Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*. 2004;23(20):4051. doi:10.1038/SJ.EMBOJ.7600385

73. Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature*. 2003;425(6956):415-419. doi:10.1038/NATURE01957
74. Gregory RI, Chendrimada TP, Shiekhattar R. MicroRNA biogenesis: isolation and characterization of the microprocessor complex. *Methods Mol Biol*. 2006;342:33-47. doi:10.1385/1-59745-123-1:33/FIGURES/1
75. Murchison EP, Hannon GJ. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr Opin Cell Biol*. 2004;16(3):223-229. doi:10.1016/J.CEB.2004.04.003
76. Lund E, Dahlberg JE. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb Symp Quant Biol*. 2006;71:59-66. doi:10.1101/SQB.2006.71.050
77. Kawahara Y, Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc Natl Acad Sci U S A*. 2012;109(9):3347-3352. doi:10.1073/PNAS.1112427109/-/DCSUPPLEMENTAL
78. Fan Z, Chen X, Chen R. Transcriptome-wide analysis of TDP-43 binding small RNAs identifies miR-NID1 (miR-8485), a novel miRNA that represses NRXN1 expression. *Genomics*. 2014;103(1):76-82. doi:10.1016/J.YGENO.2013.06.006
79. Gascon E, Lynch K, Ruan H, et al. Alterations in microRNA-124 and AMPA receptors contribute to social behavioral deficits in frontotemporal dementia. *Nat Med*. 2014;20(12):1444. doi:10.1038/NM.3717
80. Zhang Z, Almeida S, Lu Y, et al. Downregulation of MicroRNA-9 in iPSC-Derived Neurons of FTD/ALS Patients with TDP-43 Mutations. *PLoS One*. 2013;8(10):e76055. doi:10.1371/JOURNAL.PONE.0076055
81. Buratti E, De Conti L, Stuani C, Romano M, Baralle M, Baralle F. Nuclear factor TDP-43 can affect selected microRNA levels. *FEBS J*. 2010;277(10):2268-2281. doi:10.1111/J.1742-4658.2010.07643.X
82. Freischmidt A, Müller K, Ludolph AC, Weishaupt JH. Systemic dysregulation of TDP-43 binding microRNAs in amyotrophic lateral sclerosis. *Acta Neuropathol Commun*. 2013;1(1):1-9. doi:10.1186/2051-5960-1-42/TABLES/4
83. Figueroa-Romero C, Hur J, Lunn JS, et al. Expression of MicroRNAs in Human Post-mortem Amyotrophic Lateral Sclerosis Spinal Cords Provides Insight into Disease Mechanisms. *Mol Cell Neurosci*. 2016;71:34. doi:10.1016/J.MCN.2015.12.008
84. Chipman LB, Pasquinelli AE. MiRNA Targeting – Growing Beyond the Seed. *Trends Genet*. 2019;35(3):215. doi:10.1016/J.TIG.2018.12.005
85. Alami NH, Smith RB, Carrasco MA, et al. Axonal transport of TDP-43 mRNA granules in neurons is impaired by ALS-causing mutations. *Neuron*. 2014;81(3):536. doi:10.1016/J.NEURON.2013.12.018
86. Glock C, Heumüller M, Schuman EM. mRNA transport & local translation in neurons. *Curr Opin Neurobiol*. 2017;45:169-177. doi:10.1016/J.CONB.2017.05.005
87. Chu JF, Majumder P, Chatterjee B, Huang SL, Shen CKJ. TDP-43 Regulates Coupled Dendritic mRNA Transport-Translation Processes in Co-operation with FMRP and Staufen1. *Cell Rep*. 2019;29(10):3118-3133.e6. doi:10.1016/J.CELREP.2019.10.061
88. Briese M, Saal-Bauernschubert L, Lüningschrör P, et al. Loss of Tdp-43 disrupts

the axonal transcriptome of motoneurons accompanied by impaired axonal translation and mitochondria function. *Acta Neuropathol Commun*. 2020;8(1):1-16. doi:10.1186/S40478-020-00987-6/FIGURES/8
89. He C. Grand Challenge Commentary: RNA epigenetics? *Nat Chem Biol*. 2010;6(12):863-865. doi:10.1038/nchembio.482
90. Bokar JA, Rath-Shambaugh ME, Ludwiczak R, Narayan P, Rottman F. Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J Biol Chem* . 1994;269(26):17697-17704. http://www.jbc.org/content/269/26/17697.abstract.
91. Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol*. 2014;16(2):191-198. doi:10.1038/ncb2902
92. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. 2014;505(7481):117-120. doi:10.1038/nature12730
93. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 2012;149(7):1635-1646. doi:10.1016/j.cell.2012.05.003
94. Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA*. 1997;3(11):1233. /pmc/articles/PMC1369564/?report=abstract. Accessed January 25, 2022.
95. Zheng G, Dahl JA, Niu Y, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell*. 2013;49(1):18-29. doi:10.1016/j.molcel.2012.10.015
96. Horiuchi K, Kawamura T, Iwanari H, et al. Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *J Biol Chem*. 2013;288(46):33292-33302. doi:10.1074/JBC.M113.500397
97. Schwartz S, Mumbach MR, Jovanovic M, et al. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep*. 2014;8(1):284-296. doi:10.1016/j.celrep.2014.05.048
98. Jia G, Fu Y, Zhao X, et al. N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol*. 2011;7(12):885-887. doi:10.1038/nchembio.687
99. Salditt-Georgieff M, Jelinek W, Darnell JE, Furuichi Y, Morgan M, Shatkin A. Methyl labeling of HeLa cell hnRNA: a comparison with mRNA. *Cell*. 1976;7(2):227-237. doi:10.1016/0092-8674(76)90022-2
100. Barbieri I, Tzelepis K, Pandolfini L, et al. Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nat 2017 5527683*. 2017;552(7683):126-131. doi:10.1038/nature24678
101. Bertero A, Brown S, Madrigal P, et al. The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nat 2018 5557695*. 2018;555(7695):256-259. doi:10.1038/nature25784
102. Huang H, Weng H, Zhou K, et al. Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nat 2019 5677748*.

2019;567(7748):414-419. doi:10.1038/s41586-019-1016-7
103. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. 2012;485(7397):201-206. doi:10.1038/nature11112
104. Linder B, Grozhik A V, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*. 2015;12(8):767-772. doi:10.1038/nmeth.3453
105. Wei CM, Gershowitz A, Moss B. N6, O2′-dimethyladenosine a novel methylated ribonucleoside next to the 5′ terminal of animal cell and virus mRNAs. *Nat 1975 2575523*. 1975;257(5523):251-253. doi:10.1038/257251a0
106. Meyer KD. DART-seq: an antibody-free method for global m6A detection. *Nat Methods*. September 2019:1-6. doi:10.1038/s41592-019-0570-0
107. Mauer J, Luo X, Blanjoie A, et al. Reversible methylation of m6Am in the 5′ cap controls mRNA stability. *Nat 2016 5417637*. 2016;541(7637):371-375. doi:10.1038/nature21022
108. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. *N6-Methyladenosine-Dependent RNA Structural Switches Regulate RNA-Protein Interactions*. Vol 518. NIH Public Access; 2015:560. doi:10.1038/NATURE14234
109. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A*. 1974;71(10):3971-3975. http://www.ncbi.nlm.nih.gov/pubmed/4372599. Accessed September 20, 2018.
110. Perry RP, Kelley DE. Existence of methylated messenger RNA in mouse L cells. *Cell*. 1974;1(1):37-42. doi:10.1016/0092-8674(74)90153-6
111. Narayan P, Rottman FM. An in vitro system for accurate methylation of internal adenosine residues in messenger RNA. *Science (80- )*. 1988;242(4882):1159-1162. doi:10.1126/SCIENCE.3187541
112. Harper JE, Miceli SM, Roberts RJ, Manley JL. Sequence specificity of the human mRNA N6-adenosine methylase in vitro. *Nucleic Acids Res*. 1990;18(19):5735-5741. doi:10.1093/NAR/18.19.5735
113. Schapira M. Structural Chemistry of Human RNA Methyltransferases. *ACS Chem Biol*. 2016;11:50. doi:10.1021/acschembio.5b00781
114. Agarwala SD, Blitzblau HG, Hochwagen A, Fink GR. RNA methylation by the MIS complex regulates a cell fate decision in yeast. *PLoS Genet*. 2012;8(6):e1002732. doi:10.1371/journal.pgen.1002732
115. Zhong S, Li H, Bodi Z, et al. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell*. 2008;20(5):1278-1288. doi:10.1105/tpc.108.058883
116. Geula S, Moshitch-Moshkovitz S, Dominissini D, et al. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science (80- )*. 2015;347(6225):1002-1006. doi:10.1126/SCIENCE.1261417/SUPPL_FILE/1261417TABLES5.XLSX
117. Liu J, Yue Y, Han D, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol*. 2014;10(2):93-95. doi:10.1038/nchembio.1432
118. Ping X-L, Sun B-F, Wang L, et al. Mammalian WTAP is a regulatory subunit of the

RNA N6-methyladenosine methyltransferase. *Cell Res*. 2014;24(2):177-189. doi:10.1038/cr.2014.3
119. Wang P, Doxtader KA, Nam Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol Cell*. 2016;63(2):306-317. doi:10.1016/J.MOLCEL.2016.05.041
120. Wang X, Feng J, Xue Y, et al. Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex. *Nature*. 2016;534(7608):575-578. doi:10.1038/nature18298
121. Śledź P, Jinek M. Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife*. 2016;5(September). doi:10.7554/ELIFE.18434
122. Tzafrir I, Dickerman A, Brazhnik O, et al. The Arabidopsis SeedGenes Project. *Nucleic Acids Res*. 2003;31(1):90. doi:10.1093/NAR/GKG028
123. Yue Y, Liu J, Cui X, et al. VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov*. 2018;4:10. doi:10.1038/s41421-018-0019-0
124. Patil DP, Chen CK, Pickering BF, et al. m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature*. 2016;537(7620):369. doi:10.1038/NATURE19342
125. Brockdorff N, Bowness JS, Wei G. Progress toward understanding chromosome silencing by Xist RNA. *Genes Dev*. 2020;34(11-12):733-744. doi:10.1101/GAD.337196.120
126. Lindtner S, Zolotukhin AS, Uranishi H, et al. RNA-binding motif protein 15 binds to the RNA transport element RTE and provides a direct link to the NXF1 export pathway. *J Biol Chem*. 2006;281(48):36915-36928. doi:10.1074/JBC.M608745200
127. Zolotukhin AS, Uranishi H, Lindtner S, Bear J, Pavlakis GN, Felber BK. Nuclear export factor RBM15 facilitates the access of DBP5 to mRNA. *Nucleic Acids Res*. 2009;37(21):7151-7162. doi:10.1093/NAR/GKP782
128. Malovannaya A, Lanz RB, Jung SY, et al. Analysis of the Human Endogenous Coregulator Complexome. *Cell*. 2011;145(5):787. doi:10.1016/J.CELL.2011.05.006
129. Wan C, Borgeson B, Phanse S, et al. Panorama of ancient metazoan macromolecular complexes. *Nature*. 2015;525(7569):339. doi:10.1038/NATURE14877
130. Wen J, Lv R, Ma H, et al. Zc3h13 regulates nuclear RNA m6A methylation and mouse embryonic stem cell self-renewal. *Mol Cell*. 2018;69(6):1028. doi:10.1016/J.MOLCEL.2018.02.015
131. Fischer J, Koch L, Emmerling C, et al. Inactivation of the Fto gene protects from obesity. *Nat 2009 4587240*. 2009;458(7240):894-898. doi:10.1038/nature07848
132. Hess ME, Hess S, Meyer KD, et al. The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. *Nat Neurosci 2013 168*. 2013;16(8):1042-1048. doi:10.1038/nn.3449
133. Zhang C, Samanta D, Lu H, et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m6A-demethylation of NANOG mRNA. *Proc Natl Acad Sci U S A*. 2016;113(14):E2047-E2056. doi:10.1073/PNAS.1602883113/ASSET/EA9D34AD-5C88-4313-8D1E-

28480C59B621/ASSETS/GRAPHIC/PNAS.1602883113FIG06.JPEG
134. Zhang C, Zhi WI, Lu H, et al. Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. *Oncotarget*. 2016;7(40):64527-64542. doi:10.18632/ONCOTARGET.11743
135. Zhang S, Zhao BS, Zhou A, et al. m6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. *Cancer Cell*. 2017;31(4):591-606.e6. doi:10.1016/J.CCELL.2017.02.013
136. Zhu T, Roundtree IA, Wang P, et al. Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. *Cell Res 2014 2412*. 2014;24(12):1493-1496. doi:10.1038/cr.2014.152
137. Li F, Zhao D, Wu J, Shi Y. Structure of the YTH domain of human YTHDF2 in complex with an m6A mononucleotide reveals an aromatic cage for m6A recognition. *Cell Res 2014 2412*. 2014;24(12):1490-1492. doi:10.1038/cr.2014.153
138. Ries RJ, Zaccara S, Klein P, et al. m6A enhances the phase separation potential of mRNA. *Nature*. 2019;571(7765):424-428. doi:10.1038/s41586-019-1374-1
139. Wang X, Zhao BS, Roundtree IA, et al. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*. 2015;161(6):1388-1399. doi:10.1016/j.cell.2015.05.014
140. Shi H, Wang X, Lu Z, et al. YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res*. 2017;27(3):315-328. doi:10.1038/cr.2017.15
141. Kennedy EM, Bogerd HP, Kornepati AVR, et al. Posttranscriptional m(6)A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host Microbe*. 2016;19(5):675-685. doi:10.1016/J.CHOM.2016.04.002
142. Du H, Zhao Y, He J, et al. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun*. 2016;7(1):12626. doi:10.1038/ncomms12626
143. Hartmann AM, Nayler O, Schwaiger FW, Obermeier A, Stamm S. The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). *Mol Biol Cell*. 1999;10(11):3909-3926. doi:10.1091/MBC.10.11.3909/ASSET/IMAGES/LARGE/MK1191038011.JPEG
144. Xiao W, Adhikari S, Dahal U, et al. Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell*. 2016;61(4):507-519. doi:10.1016/J.MOLCEL.2016.01.012
145. Roundtree IA, Luo GZ, Zhang Z, et al. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife*. 2017;6. doi:10.7554/ELIFE.31311
146. Wojtas MN, Pandey RR, Mendel M, Homolka D, Sachidanandam R, Pillai RS. Regulation of m6A Transcripts by the 3'→5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Mol Cell*. 2017;68(2):374-387.e12. doi:10.1016/J.MOLCEL.2017.09.021
147. Bailey AS, Batista PJ, Gold RS, et al. The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. *Elife*.

2017;6. doi:10.7554/ELIFE.26116
148. Hsu PJ, Zhu Y, Ma H, et al. Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res 2017 279*. 2017;27(9):1115-1127. doi:10.1038/cr.2017.99
149. Jain D, Puno MR, Meydan C, et al. Ketu mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. *Elife*. 2018;7. doi:10.7554/ELIFE.30919
150. Xu C, Liu K, Ahmed H, Loppnau P, Schapira M, Min J. Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J Biol Chem*. 2015;290(41):24902-24913. doi:10.1074/JBC.M115.680389
151. Patil DP, Pickering BF, Jaffrey SR. Reading m6A in the Transcriptome: m6A-Binding Proteins. *Trends Cell Biol*. 2018;28(2):113-127. doi:10.1016/J.TCB.2017.10.001
152. Kretschmer J, Rao H, Hackert P, Sloan KE, Höbartner C, Bohnsack MT. The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5'-3' exoribonuclease XRN1. *RNA*. 2018;24(10):1339-1350. doi:10.1261/rna.064238.117
153. Huang H, Weng H, Sun W, et al. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol 2018 203*. 2018;20(3):285-295. doi:10.1038/s41556-018-0045-z
154. Youn JY, Dunham WH, Hong SJ, et al. High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol Cell*. 2018;69(3):517-532.e11. doi:10.1016/J.MOLCEL.2017.12.020
155. Sun L, Fazal FM, Li P, et al. RNA structure maps across mammalian cellular compartments. *Nat Struct Mol Biol*. 2019;26(4):322. doi:10.1038/S41594-019-0200-7
156. Roost C, Lynch SR, Batista PJ, Qu K, Chang HY, Kool ET. Structure and Thermodynamics of N6-Methyladenosine in RNA: A Spring-Loaded Base Modification. *J Am Chem Soc*. 2015;137(5):2107. doi:10.1021/JA513080V
157. Spitale RC, Flynn RA, Zhang QC, et al. Structural imprints in vivo decode RNA regulatory mechanisms. *Nat 2015 5197544*. 2015;519(7544):486-490. doi:10.1038/nature14263
158. Liu N, Zhou KI, Parisien M, Dai Q, Diatchenko L, Pan T. N 6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res*. 2017;45(10):6051-6063. doi:10.1093/nar/gkx141
159. Wu B, Su S, Patil DP, et al. Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1. *Nat Commun 2018 91*. 2018;9(1):1-12. doi:10.1038/s41467-017-02770-z
160. Sommer S, Lavi U, Darnell JE. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *J Mol Biol*. 1978;124(3):487-499. doi:10.1016/0022-2836(78)90183-3
161. Batista PJ, Molinie B, Wang J, et al. m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell*. 2014;15(6):707-719. doi:10.1016/J.STEM.2014.09.019
162. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of

messenger RNA stability. *Nature*. 2014;505(7481):117-120. doi:10.1038/nature12730
163. Tirumuru N, Zhao BS, Lu W, Lu Z, He C, Wu L. N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *Elife*. 2016;5(2016JULY). doi:10.7554/ELIFE.15528
164. Lu W, Tirumuru N, Gelais CS, et al. N6-Methyladenosine–binding proteins suppress HIV-1 infectivity and viral production. *J Biol Chem*. 2018;293(34):12992. doi:10.1074/JBC.RA118.004215
165. Luo Y, Na Z, Slavoff SA. P-Bodies: Composition, Properties, and Functions. *Biochemistry*. 2018;57(17):2424-2431. doi:10.1021/ACS.BIOCHEM.7B01162/ASSET/IMAGES/LARGE/BI-2017-01162X_0002.JPEG
166. Fu Y, Zhuang X. m6A-binding YTHDF proteins promote stress granule formation. *Nat Chem Biol*. 2020;16(9):955. doi:10.1038/S41589-020-0524-Y
167. Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, Parker R. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell*. 2016;164(3):487-498. doi:10.1016/J.CELL.2015.12.038
168. Anders M, Chelysheva I, Goebel I, et al. Dynamic m6A methylation facilitates mRNA triaging to stress granules. *Life Sci alliance*. 2018;1(4):e201800113. doi:10.26508/lsa.201800113
169. Du H, Zhao Y, He J, et al. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun*. 2016;7(1):12626. doi:10.1038/ncomms12626
170. Zheng D, Ezzeddine N, Chen CYA, Zhu W, He X, Shyu A Bin. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J Cell Biol*. 2008;182(1):89. doi:10.1083/JCB.200801196
171. Park OH, Ha H, Lee Y, et al. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol Cell*. March 2019. doi:10.1016/J.MOLCEL.2019.02.034
172. Boo SH, Ha H, Lee Y, Shin M-K, Lee S, Kim YK. UPF1 promotes rapid degradation of m6A-containing RNAs. *Cell Rep*. 2022;39(8):110861. doi:10.1016/J.CELREP.2022.110861/ATTACHMENT/BC3EB567-078A-4074-BFF0-6F50225914C5/MMC1.PDF
173. Haussmann IU, Bodi Z, Sanchez-Moran E, et al. m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. *Nat 2016 5407632*. 2016;540(7632):301-304. doi:10.1038/nature20577
174. Lence T, Akhtar J, Bayer M, et al. m6A modulates neuronal functions and sex determination in Drosophila. *Nat 2016 5407632*. 2016;540(7632):242-247. doi:10.1038/nature20568
175. Kan L, Grozhik A V., Vedanayagam J, et al. The m6A pathway facilitates sex determination in Drosophila. *Nat Commun 2017 81*. 2017;8(1):1-16. doi:10.1038/ncomms15737
176. Ke S, Pandya-Jones A, Saito Y, et al. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev*. 2017;31(10):990-1006. doi:10.1101/GAD.301036.117/-/DC1

177. Gokhale NS, McIntyre ABR, Mattocks MD, et al. Altered m 6 A Modification of Specific Cellular Transcripts Affects Flaviviridae Infection. *Mol Cell*. 2020;77(3):542-555.e8. doi:10.1016/J.MOLCEL.2019.11.007
178. Fish L, Navickas A, Culbertson B, et al. Nuclear TARBP2 drives oncogenic dysregulation of RNA splicing and decay. *Mol Cell*. 2019;75(5):967. doi:10.1016/J.MOLCEL.2019.06.001
179. Pendleton KE, Chen B, Liu K, et al. The U6 snRNA m6A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell*. 2017;169(5):824. doi:10.1016/J.CELL.2017.05.003
180. Edupuganti RR, Geiger S, Lindeboom RGH, et al. N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. *Nat Struct Mol Biol*. 2017;24(10):870-878. doi:10.1038/nsmb.3462
181. Zheng Q, Hou J, Zhou Y, Li Z, Cao X. The RNA helicase DDX46 inhibits innate immunity by entrapping m6A-demethylated antiviral transcripts in the nucleus. *Nat Immunol 2017 1810*. 2017;18(10):1094-1103. doi:10.1038/ni.3830
182. Chen ZH, Chen TQ, Zeng ZC, et al. Nuclear export of chimeric mRNAs depends on an lncRNA-triggered autoregulatory loop in blood malignancies. *Cell Death Dis*. 2020;11(7). doi:10.1038/S41419-020-02795-1
183. Zhang F, Kang Y, Wang M, et al. Fragile X mental retardation protein modulates the stability of its m6A-marked messenger RNA targets. *Hum Mol Genet*. 2018;27(22):3936. doi:10.1093/HMG/DDY292
184. Edens BM, Vissers C, Su J, et al. FMRP Modulates Neural Differentiation through m 6 A-Dependent mRNA Nuclear Export. *Cell Rep*. 2019;28(4):845-854.e5. doi:10.1016/J.CELREP.2019.06.072
185. Hsu PJ, Shi H, Zhu AC, et al. The RNA-binding protein FMRP facilitates the nuclear export of N6-methyladenosine–containing mRNAs. *J Biol Chem*. 2019;294(52):19889. doi:10.1074/JBC.AC119.010078
186. Slobodin B, Han R, Calderone V, et al. Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell*. 2017;169(2):326. doi:10.1016/J.CELL.2017.03.031
187. Jackson RJ, Hellen CUT, Pestova T V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol 2010 112*. 2010;11(2):113-127. doi:10.1038/nrm2838
188. Coots RA, Liu XM, Mao Y, et al. m6A Facilitates eIF4F-Independent mRNA Translation. *Mol Cell*. 2017;68(3):504. doi:10.1016/J.MOLCEL.2017.10.002
189. Choe J, Lin S, Zhang W, et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature*. 2018;561(7724):556. doi:10.1038/S41586-018-0538-8
190. Wang T, Kong S, Tao M, Ju S. The potential role of RNA N6-methyladenosine in Cancer progression. *Mol Cancer 2020 191*. 2020;19(1):1-18. doi:10.1186/S12943-020-01204-7
191. Li L, Zang L, Zhang F, et al. Fat mass and obesity-associated (FTO) protein regulates adult neurogenesis. *Hum Mol Genet*. 2017;26(13):2398-2411. doi:10.1093/HMG/DDX128
192. Chen X, Yu C, Guo M, et al. Down-Regulation of m6A mRNA Methylation Is Involved in Dopaminergic Neuronal Death. *ACS Chem Neurosci*.

2019;10(5):2355-2363. doi:10.1021/ACSCHEMNEURO.8B00657/ASSET/IMAGES/LARGE/CN-2018-00657Q_0005.JPEG
193. Quan W, Li J, Liu L, et al. Influence of N6-Methyladenosine Modification Gene HNRNPC on Cell Phenotype in Parkinson's Disease. *Parkinsons Dis*. 2021;2021. doi:10.1155/2021/9919129
194. Han M, Liu Z, Xu Y, et al. Abnormality of m6A mRNA Methylation Is Involved in Alzheimer's Disease. *Front Neurosci*. 2020;14:98. doi:10.3389/fnins.2020.00098
195. Zhao F, Xu Y, Gao S, et al. METTL3-dependent RNA m6A dysregulation contributes to neurodegeneration in Alzheimer's disease through aberrant cell cycle events. *Mol Neurodegener*. 2021;16(1):1-25. doi:10.1186/S13024-021-00484-X/FIGURES/12
196. Keller L, Xu W, Wang HX, Winblad B, Fratiglioni L, Graff C. The Obesity Related Gene, FTO, Interacts with APOE, and is Associated with Alzheimer's Disease Risk: A Prospective Cohort Study. *J Alzheimer's Dis*. 2011;23(3):461-469. doi:10.3233/JAD-2010-101068
197. Li H, Ren Y, Mao K, et al. FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tau signaling. *Biochem Biophys Res Commun*. 2018;498(1):234-239. doi:10.1016/J.BBRC.2018.02.201
198. Jiang L, Lin W, Zhang C, et al. Interaction of tau with HNRNPA2B1 and N 6-methyladenosine RNA mediates the progression of tauopathy. *Mol Cell*. 2021;81(20):4209-4227.e12. doi:10.1016/J.MOLCEL.2021.07.038
199. Koranda JL, Dore L, Shi H, et al. Mettl14 Is Essential for Epitranscriptomic Regulation of Striatal Function and Learning. *Neuron*. 2018;99(2):283-292.e5. doi:10.1016/J.NEURON.2018.06.007

## Figures

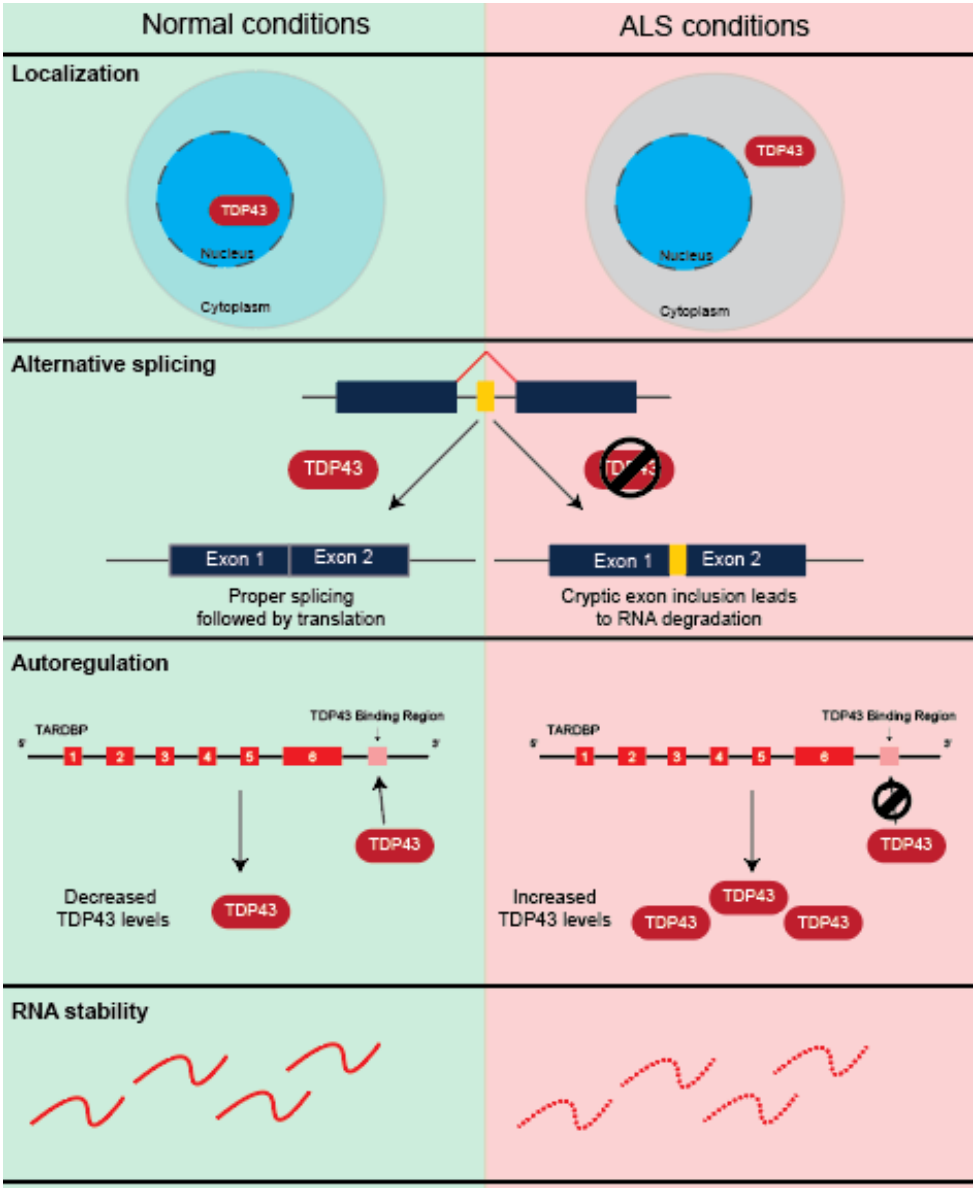



**Figure 1.1: TDP43 functions and consequences of dysregulation in ALS.** Under normal conditions, TDP43 is localized to the nucleus where it helps to maintain RNA homeostasis. In ALS conditions, TDP43 is mislocalized to the cytoplasm and loss of splicing function leads to inclusion of cryptic exons triggering early degradation of the produced transcript. TDP43 strictly regulates its protein levels to prevent TDP43-mediated toxicity by binding to the *TARDBP* 3' UTR. This binding results in alternative splicing producing a transcript degraded by nonsense-mediated decay. In ALS, mutations in *TARDBP* may prevent association of TDP43 with its transcript preventing autoregulation and allowing accumulation of TDP43 protein. TDP43 depletion in ALS also results in decreased global RNA stability, indicating the importance of TDP43 to RNA homeostasis.

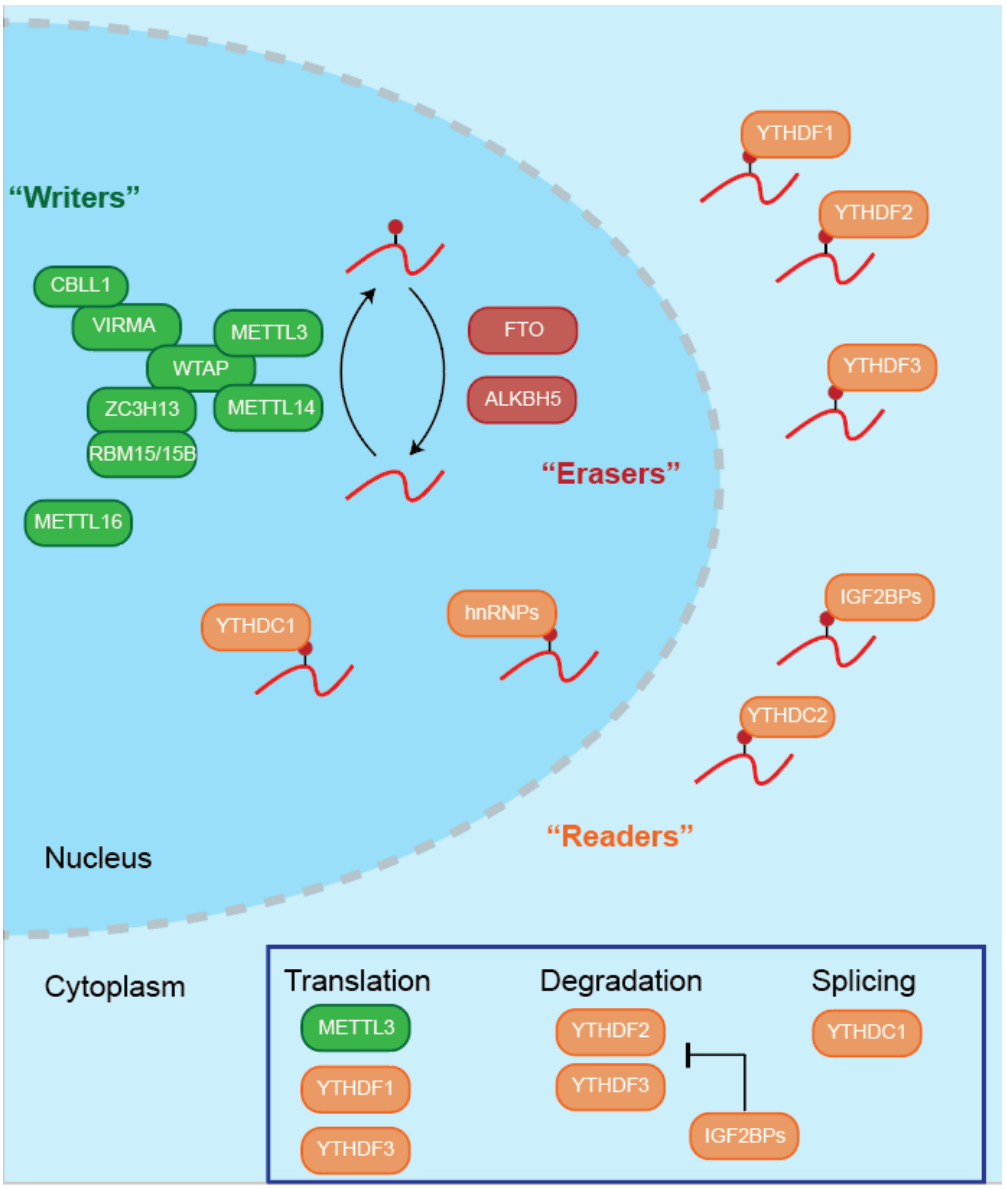



**Figure 1.2: m6A RNA methylation pathway.** m6A methylation (pin) is deposited onto RNA (red lines) by methyltransferases ("writers") and can be removed or prevent the addition of by demethylases ("erasers"). RNA binding proteins that recognize m6A methylation on RNAs ("readers") influence the fate of RNA by increasing the likelihood of that RNA to be translated, degraded, or spliced.

# Chapter 2: RNA Methylation Influences TDP43 Binding and Disease Pathogenesis in Models of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia[1]

## 2.1 Abstract

Methylation of RNA at the N6 position of adenosine (m6A) is one of the most common RNA modifications, impacting RNA stability as well as its transport and translation. Previous studies uncovered RNA destabilization in models of amyotrophic lateral sclerosis (ALS), in association with accumulation of the RNA-binding protein TDP43, which is itself mislocalized from the nucleus in >95% of those with ALS. Here, we show that TDP43 recognizes m6A-modified RNA, and that RNA methylation is critical for both TDP43 binding and autoregulation. We also observed extensive hypermethylation of coding and non-coding transcripts in ALS spinal cord, many of which overlap with methylated TDP43 target RNAs. Emphasizing the importance of m6A for TDP43 binding and function, we identified several m6A factors that enhance or suppress TDP43-mediated toxicity via a single-cell CRISPR/Cas9 candidate-based screen in primary neurons. The most promising genetic modifier among these—the canonical m6A reader YTHDF2—accumulated within spinal motor neurons in ALS postmortem sections, and its knockdown prolonged the survival of human neurons carrying ALS-associated mutations.

---

[1] This chapter is adapted from McMillan M, Gomez N, Bekier M, Li X, Tank EM, Barmada SJ. RNA methylation influences TDP43 binding and disease pathogenesis in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Under review Molecular Cell, Cell Reports*. Published online b*ioRxiv*. 2022. doi:10.1101/2022.04.03.486880

Collectively, these data show that m6A modifications modulate RNA binding by TDP43, and that m6A is pivotal for TDP43-related neurodegeneration in ALS.

## 2.2 Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease resulting in the death of upper and lower motor neurons[1]. Limited therapeutic options exist for this condition, and the underlying pathological mechanisms remain unclear. Considerable variability in clinical, biochemical, and genetic features also complicate identification of new therapeutic targets. Despite this, over 95% of ALS cases exhibit cytoplasmic inclusions of the RNA binding protein TDP43 (TAR binding protein of 43 kDa), and mutations in *TARDBP*, the gene encoding TDP43, result in familial disease in 2-5% of individuals[5,6,59,198,199]. These observations imply that strategies targeting TDP43 may be relevant to the large majority of those with ALS.

TDP43 is a predominantly nuclear RNA binding protein critical for several aspects of RNA processing, including RNA splicing, transport, translation, and stability. Consistent with this, cytoplasmic mislocalization and nuclear clearing of TDP43 in ALS are closely associated with RNA missplicing, abnormal global and local mRNA translation, and widespread RNA instability[21,55,56,59]. We previously found that TDP43 deposition primarily destabilizes families of mRNAs encoding ribosomal proteins and oxidative phosphorylation enzymes[59]. Notably, these same mRNAs are upregulated in neurons lacking the RNA methyltransferase-like protein 14 (METTL14)[200]. This enzyme acts as a part of a “writer” complex that methylates RNA at the 6$^{th}$ position nitrogen ($N^6$-methyladenosine methylation, or m6A)[91,101]. m6A marks can be removed by demethylases (“erasers”) and/or recognized by a class of RNA binding proteins

("readers") that, like TDP43, function in RNA splicing, transport, translation and stability[90,158]. In addition, TDP43 binds heavily methylated RNAs, and physically interacts with both writers and readers[201–204]. Together, these findings suggest that TDP43 may recognize m6A-modified RNA, raising the possibility that mislocalization and aggregation of TDP43 in ALS may preferentially affect RNA substrates carrying m6A marks.

Here, we answer the question of whether TDP43 binds m6A-modified RNA through several orthologous approaches. We show not only that the majority of TDP43 substrates carry m6A modifications, but also that the m6A reader protein YTHDF2 facilitates TDP43-related toxicity in rodent and human neuron models of ALS. Supporting the connection between TDP43 pathology in ALS and m6A RNA, we detected extensive RNA hypermethylation in postmortem spinal cord tissue from sporadic ALS patients compared to controls. These data highlight a fundamental link between m6A RNA modifications and ALS pathogenesis, potentially mediated by TDP43-dependent misprocessing of m6A-modified RNA.

## 2.3 Results

### *2.3.1 TDP43 binds m6A-modified RNA*

YTH domains, common to m6A reader proteins, exhibit selective recognition of m6A-modified RNA, but previous evidence indicates that other functional domains such as RNA recognition motifs (RRMs) present in TDP43 and other members of the hnRNP family can also bind m6A-modified RNA[101,160]. Specifically, hnRNP-C[106] and hnRNP-A2/B1[205] are capable of recognizing m6A-modified RNA despite not having YTH domains, and the presence of m6A modifications enhances their affinity for RNA targets. Given this, we asked whether TDP43, which shows high affinity for UG rich sequences[13,15,206], also

recognizes m6A-modified RNA. First, we measured the UG density surrounding experimentally verified m6A sites, determined by m6A antibody cross-linking induced methylation (CIMS; Fig. 2.1A) and truncation (CITS; Fig. 2.1B)[102]. UG density is significantly greater immediately upstream of m6A sites, whereas random segments of genes of non-methylated genes showed no apparent change in UG density surrounding the m6A site (Fig. 2.1A, B). These data indicate a possible connection between UG-rich TDP43 recognition motifs and m6A sites. To pursue this concept further and empirically determine if TDP43 recognizes m6A-modifed RNA in mammalian cells, we overexpressed TDP43 fused with HaloTag, a multifunctional adapter protein that facilitates TDP43 isolation by immunoaffinity purification[207], in HEK293T cells. As a positive control, we also overexpressed a fusion of HaloTag and YTHDF2, a verified m6A reader protein[160]. We then isolated each protein by immunoaffinity purification, collected bound RNA and assessed total and m6A RNA by dot blot (Fig. 2.1C). TDP43-HaloTag and YTHDF2-HaloTag, but not HaloTag by itself, recognized and pulled down m6A RNA (Fig. 2.1D), showing that TDP43 and YTHDF2 are both capable of binding m6A RNA.

Because protein overexpression can introduce non-specific and potentially non-physiological interactions, we engineered a line of HEK293T cells in which endogenous TDP43 was labeled at the N-terminus with HaloTag (Fig. 2.1E). Correct insertion of HaloTag into the *TARDBP* locus was verified by Sanger sequencing, and analysis of HaloTag-TDP43 function using the *CFTR* splicing reporter uncovered no loss-of-function effects associated with the fusion (Sup. Fig. 2.1A). In addition, live-cell labeling of HaloTag-TDP43 HEK293T cells with JF646, a far-red cell-permeable dye that binds covalently to HaloTag, showed a strong overlap with TDP43 localization, as judged by

immunostaining using TDP43 antibodies (Fig. 2.1F). We then utilized these cells to determine if endogenous HaloTag-TDP43 is capable of recognizing m6A modified RNA. As before, we detected m6A modified RNA via dot blot among transcripts pulled down by HaloTag-TDP43 (Fig. 2.1G and Sup. Fig. 2.1B-C), confirming that endogenously expressed TDP43 is capable of binding m6A modified RNA.

### *2.3.2 TDP43 substrates are enriched in m6A modifications*

To explicitly define the identity, number, and location of m6A modifications in TDP43 substrates, we took advantage of DART-seq (deamination adjacent to RNA modification targets, followed by next generation RNA sequencing)[104] (Fig. 2A). This method highlights specific m6A modifications without the need for m6A antibodies, which may have limited sensitivity and specificity in discriminating m6A from other RNA modifications (i.e., m6Am). Briefly, DART-seq involves the overexpression of a chimeric fusion protein consisting of the m6A-binding YTH domain, and the deaminating enzyme APOBEC1. As illustrated previously[91,101,102,104], m6A modifications occur in the context of a highly conserved motif (DRACH, where D can be A, G or T; R is A or G; and H is A, C or T). Within this motif, the methylated adenosine (A) residue is followed by an obligate cytosine (C). Recognition of m6A by the APOBEC1-YTH fusion results in deamination of the C immediately 3' to the modified A residue, creating a uracil (U) base that is read as a thymine (T) during sequencing. Thus, C-T transitions occurring within a DRACH motif immediately adjacent to an A signify m6A modifications. In combination with immunoaffinity purification of endogenous HaloTag-TDP43 from the HEK293T cells we created, DART-seq enables us to define not only the RNA substrates of native HaloTag-TDP43, but also the location and m6A modification status of each substrate.

For these experiments, we transfected HaloTag-TDP43 HEK293T cells with APOBEC1-YTH or a mutated variant of APOBEC1-YTH that is unable to bind m6A modified RNA (APOBEC1-YTHmut). To control for non-specific RNA binding by HaloTag, we also transfected separate cultures of unmodified HEK293T cells with HaloTag alone. We then isolated HaloTag and HaloTag-TDP43 by immunoaffinity purification, sequenced the RNA that was pulled down in each case, and compared the resulting data to a reference database to identify base-pair transitions (Fig. 2.2A). As expected based on the deaminase activity of APOBEC1, C-T and G-A transitions—and to a lesser extent antisense T-C and A-G transitions—were strongly enriched in APOBEC1-YTH expressing cells in comparison to those expressing APOBEC1-YTHmut or HaloTag alone (Fig. 2.2B, C). These transitions occurred both within and outside of canonical DRACH motifs (Fig. 2.2D), but to increase specificity and confidence in the detection of verified m6A sites, we limited further analyses to transitions that arose in the context of a DRACH motif. In doing so, we noted profound enrichment for high-confidence m6A sites in cells expressing APOBEC1-YTH, compared to those expressing APOBEC1-YTHmut and HaloTag (Fig. 2.2E, F). The majority of detected m6A sites fall within the coding sequence (CDS) of HaloTag-TDP43 target RNAs, rather than the untranslated regions (5’ or 3’UTRs) or intronic segments (Sup. Fig. 2.2A, B). This pattern is consistent with the previously observed exclusion of m6A sites from introns, but differs from the expected concentration of m6A modifications in the 3’UTR immediately downstream of the stop codon[91,102,104]. We suspect that the unusual distribution of m6A sites likely reflects the differences between HaloTag-TDP43 substrates investigated here, and total mRNA used in prior studies.

Emphasizing the sensitivity of this approach for capturing TDP43 target RNAs, we identified 2,256 transcripts that were significantly enriched by HaloTag-TDP43 pulldown in APOBEC1-YTH and APOBEC1-YTHmut expressing cells, compared to cells expressing HaloTag alone (Fig. 2.2G). In keeping with our previous data suggesting that TDP43 recognizes m6A-modified RNA, the majority of HaloTag-TDP43 target RNAs were methylated (Fig. 2.2H), and also displayed a higher degree of methylation in comparison to non-targets. (Fig. 2.2I). To ensure that our results are not confounded by potential non-specific RNA-protein interactions of HaloTag or the immunoaffinity purification itself, we compared our HaloTag-TDP43 substrates to a published dataset of TDP43 targets in HEK293 cells identified by CLIP-seq[208]. In the process, we identified 1822 high-confidence TDP43 substrates ($p= 1.5x10^{-117}$ for the overlap, hypergeometric test), over 90% (1699) of which were m6A-modified (Fig. 2.2J, K). By gene ontology (GO), methylated TDP43 target RNAs were highly enriched for components of the nuclear pore complex (Sup. Fig. 2.3A), a structure with intricate ties to ALS pathogenesis. Furthermore, protein-protein interaction (PPI) network prediction using STRING[209] indicated strong enrichment for several additional disease-associated pathways, including apoptosis and p53 signaling[210], the ribosome[59], long-term potentiation[211], VEGF signaling[212] and RNA transport[213,214] (Sup. Fig. 2.3B-F). In comparison, only 62% of non-targets were methylated ($p < 1x10^{-5}$ for the comparison with TDP43 targets; chi-square test). These results not only confirm that TDP43 recognizes m6A modified RNA, but also indicate a possible preference of TDP43 for methylated transcripts.

### *2.3.3 RNA methylation modulates TDP43 binding and autoregulation*

To explore the functional impact of RNA methylation on TDP43 binding and function, we re-examined the DART-seq results, concentrating on the relationship between identified m6A sites and UG-rich TDP43 recognition motifs. In doing so, we observed a striking overlap between m6A modifications and UG-rich sequences primarily within the 3'UTR of HaloTag-TDP43 substrates (Sup. Fig. 2.2C); indeed, 74% of m6A sites located within 20 nt of a TDP43 motif are found within the 3'UTR of HaloTag-TDP43 targets (Sup. Fig. 2.2D). This subset of 237 transcripts was highly enriched for genes whose expression is regulated by TDP43 (Sup. Fig. 2.2E), emphasizing the functional relevance of m6A sites for TDP43-dependent regulation.

We next focused on *TARDBP*, a HaloTag-TDP43 substrate RNA that showed prominent methylation based on our DART-seq experiments. The *TARDBP* transcript, which encodes TDP43 itself, exhibited a clear C-T transition within the context of a 3'UTR DRACH motif (Fig. 2.3A). We focused on this m6A site for three reasons: First, it is located within the TDP43 binding region (TBR)[215], a section of the 3'UTR that is both necessary and sufficient for TDP43 recognition. Second, the m6A site is immediately adjacent to a UG-rich stretch resembling the consensus GUGUGU motif commonly recognized by TDP43[20,37]. Third, the m6A site lies within a 34 nucleotide stretch (denoted CLIP34nt) that binds TDP43 *in vitro* and *in cellulo*[49] (Fig. 2.3A).

To probe the importance of this m6A modification for TDP43 binding and function, we utilized a minigene reporter in which *TARDBP* exon 6 and a portion of the 3'UTR containing the TBR is fused to mCherry (Fig. 2.3B; mCherry-TBR)[21,216,217]. We then applied site-directed mutagenesis to change the methylated adenosine residue shown in Fig. 2.3A to a guanosine, thereby blocking methylation of the mutated reporter (mCherry-

mTBR). Both reporters were then expressed in HaloTag-TDP43 HEK293T cells or HEK293T cells transfected with HaloTag alone. Endogenous HaloTag-TDP43 was isolated by immunoaffinity purification, and bound RNA separated and assessed by qRT-PCR (Fig. 2.3C). HaloTag-TDP43 efficiently pulled down the mCherry-TBR reporter but not the m6A-deficient mCherry-mTBR reporter despite no change in reporter input levels (Sup. Fig. 2.4A), demonstrating that the m6A modification is required for recognition by TDP43.

Two possibilities may account for the apparent effect of m6A modifications on RNA recognition by TDP43: a direct influence on binding, resulting from an increase in affinity of TDP43 for m6A *vs.* unmodified RNA; or an indirect effect arising from enhanced accessibility of the TDP43 binding motif in m6A-modified RNA. Several observations argue for an indirect rather than a direct effect of m6A modifications on TDP43 binding: First, TDP43 typically recognizes UG-rich motifs rather than the DRACH sequence associated with m6A[20,37]. Second, TDP43 exhibits canonical RRM motifs rather than the more m6A-specific YTH domains found in most direct m6A reader proteins[218]. Third, hnRNP-C[106], a heterogeneous ribonucleoprotein that is structurally and functionally related to TDP43, is an indirect m6A binding protein. To examine these possibilities in more detail, we created short (14 nt) synthetic RNA probes corresponding to the TBR with or without methylation of the predicted m6A site and performed electromobility shift assays (EMSAs) with recombinant TDP43, enabling measurements of probe binding *in vitro* (Sup. Fig. 2.4B, C). In contrast to results from RNA-immunoprecipitation studies (Fig. 2.3B), methylation of the TBR probe had little effect on TDP43 binding. In fact, we observed a slight (~2-fold) reduction in binding affinity for the m6A-modified probe in

comparison to the unmodified probe. This discrepancy could arise from differences in secondary structure between RNA transcripts present in cells and the short 14 nt probe used in these experiments, or the absence of key cellular components absent from the *in vitro* assay. Regardless, these studies suggest that m6A modifications indirectly affect TDP43 binding, perhaps by enabling greater access to UG-rich motifs buried within the TBR.

One consequence of TDP43 binding to the TBR is downregulation of TDP43 expression—this autoregulatory feedback loop is crucial for maintaining TDP43 homeostasis[20,50,199,219–221]. The mCherry-TBR reporter includes the majority of exon 6 and the *TARDBP* 3'UTR, facilitating investigations of TDP43 autoregulation in response to engineered mutations as well as genetic modulators of TDP43 function and toxicity[21,216,217]. Binding of TDP43 to the TBR in the mCherry-TBR reporter results in 3'UTR splicing and a reduction of mCherry fluorescence (Fig. 2.3D), likely due to a combination of nuclear retention of the spliced transcript and nonsense mediated mRNA decay (NMD)[220]. Consistent with this, TDP43-EGFP overexpression reduces mCherry fluorescence in primary neurons transfected with the mCherry-TBR reporter (Fig. 2.3E, F; Sup. Fig. 2.4D). In contrast, the fluorescence intensity of the m6A-deficient mCherry-mTBR reporter was unaffected by TDP43-EGFP expression (Fig. 2.3G; Sup. Fig. 2.4D). Thus, not only is the TBR m6A site required for recognition of the mCherry-TBR reporter, but it also affects the efficiency of TDP43 autoregulation.

### *2.3.4 RNA hypermethylation in ALS spinal cord*

Based on data in Figs. 2.1 and 2.2 suggesting that TDP43 binds methylated RNA, and the prominent effect of RNA methylation on TDP43 autoregulation (Fig. 2.3), we

questioned whether RNA methylation may be disrupted in sporadic ALS (sALS), a disorder characterized by nuclear clearance and cytoplasmic accumulation of TDP43[5,6,198]. To answer this question, we obtained fresh frozen spinal cord from 4 sALS patients and 3 age-matched controls (Table 2.1), isolated RNA from these samples, and quantitatively assessed transcript methylation using an m6A array (Fig. 2.4A).

Principal component analysis (PCA; Fig. 2.4B) and hierarchical clustering (Fig. 2.4C) demonstrated distinct patterns of mRNA methylation in sALS vs. control sections. These studies also suggested widespread mRNA hypermethylation in sALS patient samples, with 37% of assayed m6A sites in sALS displaying increased methylation (Fig. 2.4D) whereas <1% of transcripts were hypomethylated compared to controls in sALS spinal cord. We detected a similar pattern of hypermethylation for lncRNA (Fig. 2.4E, F), indicating a broad phenomenon not limited to protein-coding mRNAs.

To determine which of the hypermethylated transcripts in ALS are m6A-modified TDP43 substrates, we examined the overlap between the DART-seq results from HEK293T cells (Fig. 2.2), and hypermethylated transcripts from sALS spinal cord (Fig. 2.4G). Within the 2184 hypermethylated transcripts that are detectable in both HEK293T cells and spinal cord, 2034 or 93% were also identified as m6A-modified RNAs by DART-seq, indicating excellent agreement between the two approaches (p~0, hypergeometric test). These transcripts were enriched for protein kinases and RNA binding proteins, several of which are also associated with ALS (*CSNK1E, TIA1, hnRNPAB, ANXA11*), other neurodegenerative diseases (*PARK7, WDR45*), basic processes such as stress granule formation (*UBAP2L*), and nucleocytoplasmic transport (*NXF1, KPNB1*). By cross-referencing these transcripts with m6A-modified RNAs pulled-down by HaloTag-TDP43

(Fig. 2.2), we identified 302 hypermethylated TDP43-target RNAs (Fig. 2.4G) — these transcripts were likewise enriched in kinases and RNA binding proteins, and included several factors linked with ALS pathogenesis (*TP53, UCHL1*) or RNA methylation itself (*METTL3, HNRNPC*). Remarkably, the 302 hypermethylated TDP43 targets, as well as the larger set of 2034 hypermethylated transcripts in sALS spinal cord, displayed strong enrichment for RNAs whose expression is regulated by TDP43 (Fig. 2.4H). Thus, a broad range of transcripts are hypermethylated in sALS spinal cord, many of which are functionally regulated by TDP43 and overlap with m6A-modified TDP43 target RNAs highlighted by DART-seq.

As confirmation for the array, and to examine the distribution of RNA hypermethylation in sALS, we immunostained control and sALS patient spinal cord using antibodies against m6A (Fig. 2.4I). Moderate cytoplasmic staining for m6A RNA was detected in large neurons located within the anterior horn of control spinal cord. In comparison, we observed pronounced and often punctate m6A staining in anterior horn neurons from sALS sections. Indeed, m6A immunoreactivity was approximately 1.5-fold greater in sALS spinal neurons compared to controls (Fig. 2.4J). These data provide independent verification of the epitranscriptomic array (Fig. 2.4C-F) demonstrating RNA hypermethylation in sALS spinal cord, and further suggest that m6A-modified RNA accumulates predominantly within spinal motor neurons in sALS.

### *2.3.5 YTHDF2 knockout mitigates TDP43-related neurotoxicity*

Considering TDP43's ability to recognize m6A-modified RNA, and the substantial RNA hypermethylation noted in sALS spinal cord, we questioned whether RNA methylation or the factors acting on m6A-modified RNA are involved in TDP43-mediated

neurotoxicity. We therefore designed a platform that would enable us to readily screen m6A-related factors using CRISPR/Cas9 in a neuron model of ALS/FTD due to TDP43 accumulation[21,216,222,223]. As proof of principle, single-guide (sg)RNAs directed against the neuronal transcription factor NeuN or lacZ were cloned into a vector that also encodes enhanced green fluorescent protein (EGFP) and Cas9 nuclease. Rodent primary mixed cortical neurons were transfected with this vector, and after 5d, cells were fixed and immunostained for NeuN. Confirming the utility of this platform for single-cell gene knockouts, we observed a marked reduction in NeuN immunoreactivity only in transfected cells marked by EGFP fluorescence, and only in neurons that received sgRNAs against NeuN (Fig. 2.5A, B).

We then designed a small-scale screen targeting the major m6A writers (METTL3, METTL14, WTAP), erasers (FTO, ALKBH5), and readers (YTHDF1, YTHDF2, YTHDF3) (Fig. 2.5C) in a neuronal model of TDP43-related disease[21,216,222,223]. sgRNAs targeting each factor were cloned into the EGFP-Cas9 vector described above and transfected into rodent primary mixed cortical neurons along with the red fluorescent protein mApple or TDP43 fused with mApple (TDP43-mApple; Fig. 2.5D). Overexpression of TDP43 recapitulates key aspects of ALS/FTD pathophysiology in neurons, including progressive neurodegeneration accompanied by TDP43 mislocalization and aggregation[32]. To track TDP43-dependent neuron loss over time, we employed automated longitudinal microscopy[224,225]. Here, hundreds of transfected neurons are simultaneously imaged at regular 24h intervals. Image segmentation algorithms detect neurons based on their unique morphology and determine their time of death through characteristic changes

including soma rounding, neurite retraction and fragmentation, and cellular blebbing (Fig. 2.5D).

Transient transfection in this manner results in a wide range of expression levels, ranging from 1-7-fold endogenous TDP43 (Sup. Fig. 2.5A). As in previous studies, we observed dose-dependent TDP43-mediated toxicity in primary neurons, with the greatest risk of death observed in neurons with the highest TDP43 expression (Sup. Fig. 2.5B-C). Therefore, to maximize our ability to identify genetic modulators of TDP43 toxicity we focused on neurons within the lowest TDP43 expression (<2-fold endogenous levels), thereby avoiding supraphysiological TDP43 levels that may mask modifier effects.

In cells expressing non-targeting (NT) sgRNA, TDP43-mApple overexpression resulted in a >3-fold increase in the risk of death compared to neurons transfected with mApple alone (Fig. 2.5E**;** hazard ratio (HR)= 3.4; p= $2x10^{-16}$, Cox proportional hazards analysis). As a positive control, we co-transfected neurons with sgRNAs targeting Atxn2, one of the strongest genetic modifiers of TDP43-dependent toxicity[226–228]. Co-expression of Atxn2 sgRNA significantly suppressed TDP43-mediated toxicity (HR= 0.80; ***p= $5.81x10^{-4}$, Cox proportional hazards analysis), validating the use of this system for identifying disease modifiers. We then assessed whether knockout of the m6A writers, erasers or readers listed above are capable of modulating neuron loss due to TDP43 overexpression (Fig. 2.5F). In control neurons expressing mApple alone, we observed baseline increases in the risk of death upon knockdown of several factors (Sup. Fig. 2.7), suggesting that many m6A-related components are essential. Among these factors, knockout of Wtap (HR= 1.3; *p= 0.02), Alkbh5 (HR= 2.03; ***p= $2.65x10^{-11}$), Ythdf1 (HR= 1.23; *p= 0.01), and Ythdf3 (HR= 1.23; *p= 0.03) significantly enhanced TDP43-related

toxicity (Fig. 2.5G and Sup. Figs. 2.6,2.7). Only Ythdf2 knockout (HR= 0.71; ***p= $6.2x10^{-6}$) reduced TDP43-mediated toxicity (Fig. 2.5H), suggesting that TDP43 may act in concert with YTHDF2 to elicit neurodegeneration.

In a complementary set of experiments, we overexpressed a subset of m6A-related factors in rodent primary neurons and followed neuronal survival by automated longitudinal microscopy. Analogous to TDP43, overexpression of YTHDF2 resulted in a significant increase in the risk of death compared to the negative control (Fig. 2.5I). Conversely, overexpression of METTL3 and METTL14 had little effect on their own in primary neurons (Fig. 2.5J). To determine if METTL3/METTL14 might act synergistically with TDP43, we expressed TDP43-mApple at low concentrations, resulting in more subtle toxicity than in previous experiments. In doing so, we found that METTL3/METTL14 overexpression significantly enhanced TDP43-related toxicity (Fig. 2.5J). Considering the ability of TDP43 to bind m6A-modified RNA, and the increase in RNA methylation stimulated by METTL3/METTL14 overexpression[115], these data imply that RNA hypermethylation facilitates TDP43-related neurodegeneration in ALS/FTD models.

### *2.3.6 YTHDF2 in sALS spinal cord and human iPSC-derived neurons*

We detected relative RNA hypermethylation in sALS spinal cord samples by m6A array (Fig. 2.4) and found that genetic ablation of the m6A reader YTHDF2 reduced TDP43-dependent toxicity in a neuronal model of disease (Fig. 2.5H). Furthermore, YTHDF2 overexpression was sufficient to elicit toxicity in rodent primary neurons (Fig. 2.5I). To determine if YTHDF2, like TDP43, is mislocalized and/or accumulates in ALS, we performed immunohistochemistry for YTHDF2 in human post-mortem samples from sALS and control spinal cord (Fig. 2.6A, B) and frontal cortex (Sup. Fig. 2.8). In controls,

YTHDF2 displayed the expected uniform, cytoplasmic distribution within neurons and other cell types. Not only was YTHDF2 staining significantly more intense within sALS spinal neurons, but we also noted punctate accumulations of YTHDF2 within many of these cells (Fig. 2.6A, B). Strong YTHDF2 staining was also noted within layer IV-V neurons from ALS frontal cortex (Sup. Fig. 2.8), but puncta were not as evident in these cells as in spinal motor neurons. These results show that YTHDF2 accumulates in association with RNA hypermethylation in sALS spinal neurons, potentially contributing to TDP43-mediated toxicity and disease pathogenesis.

The overabundance of YTHDF2 in sALS spinal cord, together with data from Fig. 2.5 suggesting neuroprotection upon Ythdf2 knockout in primary neurons, imply that YTHDF2 could be a therapeutic target for sALS. To pursue this further, we examined the effect of *YTHDF2* knockdown in human neuron disease models. The ALS-associated TDP43(M337V) mutation was introduced into the native *TARDBP* locus of control human induced pluripotent stem cells (iPSCs) by CRISPR/Cas9[229]. Simultaneously, endogenous TDP43 was labeled at the C-terminus with Dendra2, facilitating identification and selection of clonal cell populations harboring the M337V mutation. We also generated isogenic WT iPSCs in which native TDP43 was fused to Dendra2 without introduction of a pathogenic mutation (Fig. 2.6C). iPSCs were then differentiated into forebrain-like neurons (iNeurons) through the induced expression of master transcription factors Ngn1-2, as described previously[217,229,230]. Neurons were tracked by longitudinal microscopy over a 10d period, and their survival assessed by semi-automated tracking software developed specifically for these purposes (Fig. 2.6D). We detected a significant increase in the risk of death for TDP43(M337V) iNeurons in comparison to controls (Fig. 2.6E, †p=

$8.42x10^{-12}$, HR= 6.25; Cox proportional hazards analysis) after sustained growth in the same culture after differentiation (see methods). As in our primary neuron disease model, *YTHDF2* knockdown via lentiviral shRNA delivery substantially extended the survival of TDP43(M337V) iNeurons without adversely affecting WT iNeurons (Fig. 2.6E, ***p= $4.82x10^{-9}$, HR= 0.32; #p= 0.08, HR= 1.84; Cox proportional hazards analysis).

We also evaluated the effect of *YTHDF2* knockdown in 3 separate lines of human iNeurons carrying the *C9ORF72* hexanucleotide expansion, the most prevalent cause of familial ALS and FTD in Northern Europe and North America[231–233]. *C9ORF72* mutant iNeurons displayed a ~3-fold elevation in the risk of death upon neurotrophic factor withdrawal compared to control neurons (Fig. 2.6F, †p= $1.42x10^{-11}$, HR= 2.86; Cox proportional hazards analysis). As in TDP43(M337V) iNeurons, shRNA-mediated knockdown of *YTHDF2* significantly prolonged the survival of *C9ORF72* mutant iNeurons but had no effect on non-disease (ND) iNeurons (***p= $1.42x10^{-16}$, HR= 0.32; Cox proportional hazards analysis). Together with our data from rodent primary neurons (Fig. 2.5), these findings emphasize the neuroprotective potential of *YTHDF2* knockdown in ALS/FTD disease models featuring TDP43 pathology.

## 2.4 Discussion

Here, we show not only that TDP43 recognizes m6A-modified RNA, but also that the majority of TDP43 substrates exhibit m6A marks. Sequence analysis demonstrated a spatial correlation between 3'UTR m6A modifications and TDP43 recognition motifs, and further showed that methylation strongly influences the binding of TDP43 to its RNA targets as well as TDP43 autoregulation. In ALS, where cytoplasmic mislocalization and aggregation of TDP43 are signature pathologic changes, we also observed widespread

RNA hypermethylation compared to disease controls. Consistent with a primary role for RNA hypermethylation in ALS pathogenesis, knockout or knockdown of the m6A writer YTHDF2 mitigated toxicity in primary rodent and human iPSC-derived neuron models of ALS and FTD, while overexpression of the m6A writers METTL3 and METTL14 exacerbated TDP43-mediated neuron loss. Together, these findings underscore the importance of m6A modifications for RNA binding by TDP43 and emphasize the potential contribution of TDP43's actions on m6A-modified RNA to the development of ALS and FTD.

Our study builds on several lines of evidence hinting at a connection between TDP43 and m6A-modified RNA. m6A writers (METTL3, METTL14, WTAP) and erasers (FTO, ALKBH5) primarily localize to nuclear speckles[92,116], membraneless organelles that facilitate peri-transcriptional RNA processing and splicing. TDP43 is also concentrated within these structures[234], and interacts with several m6A writer and reader proteins, including METTL3, YTHDF1, YTHDF2 and hnRNPC[95,201,235]. Furthermore, TDP43 was among a series of proteins that exhibited preferential binding to m6A-modified bait RNAs[178]. Supporting this, we found that both overexpressed and endogenously labeled TDP43 in HEK293T cells were capable of pulling down m6A modified RNA. Combining affinity purification of native HaloTag-TDP43 with antibody-free detection of m6A modifications through DART-seq, we were able to clarify the location and number of m6A sites within TDP43 substrate RNAs. Importantly, DART-seq is not limited by the specificity or sensitivity of antibodies for m6A modifications, instead relying on APOBEC1-YTH induced C-T transitions to indicate possible m6A sites. In contrast to the previously noted enrichment of m6A marks within the proximal 3'UTR of transcripts[91,101,102,104], TDP43

targets displayed a broad distribution of m6A modifications across the CDS and 3'UTR. It is unclear whether this pattern is unique to TDP43 substrate RNAs, or whether a similar distribution of m6A marks would be observed in targets of related RNA binding proteins. We observed few m6A sites within the introns of TDP43 target RNAs, consistent with the relative lack of intronic m6A modifications noted in previous studies[174,236,237]. Given the relative concentration of TDP43 binding sites within introns and 3'UTRs[20,37], and our finding that m6A modifications cluster within the CDS and 3'UTR of TDP43 target RNAs, it stands to reason that the overlap between TDP43 binding sites and m6A modifications is greatest within the 3'UTR of these transcripts. We anticipate that the expression and/or stability of this subset of TDP43 target RNAs is strongly influenced by conditions that enhance or reduce m6A modifications.

To assess the consequences of such m6A modifications, we focused on the *TARDBP* transcript itself. Previous studies as well as our own DART-seq experiments identified key m6A modifications located within the TBR, a region of the *TARBDP* transcript that is essential for recognition and autoregulation by TDP43 itself[49,50]. In keeping with the association between m6A modifications and UG-rich domains recognized by TDP43 (Fig. 2.1), these m6A marks are located immediately adjacent to TDP43 binding motifs within the TBR. Mutagenesis of the methylated A within a reporter containing the TBR led to a reduction in TDP43 binding and ineffective autoregulation, suggesting that m6A modifications are crucial for TDP43 recognition and events downstream of binding. Even so, *in vitro* electromobility shift assays involving recombinant TDP43 and a short (14 nt) m6A-modified probe failed to show methylation-dependent increases in TDP43 binding, implying that m6A modifications indirectly

enhance RNA recognition by TDP43. Similar relationships have been observed for hnRNP-C[106] and IGF2BP3[153], suggesting that changes in RNA secondary or tertiary structure upon methylation can promote recognition of otherwise buried motifs by TDP43 and other RNA binding proteins.

In prior work, we observed extensive RNA destabilization in human iPSCs overexpressing TDP43[59]. Transcripts encoding components of the ribosome and oxidative phosphorylation pathways were most heavily represented among TDP43-destabilized RNAs, a pattern that was mirrored in iPSCs carrying ALS/FTD-associated *C9ORF72* mutations. Notably, these same families of transcripts are upregulated upon targeted deletion of the m6A writer METTL14[200], indicating that they are regulated by m6A modification. Given these observations, together with the degree of RNA hypermethylation we observed in sALS spinal cord and the ability of TDP43 to recognize m6A-modified RNA, we suspect that TDP43 mislocalization and accumulation may lead to RNA destabilization in ALS through an m6A-dependent mechanism. Although this hypothesis requires further investigation, as an example we explored the potential for m6A modifications to influence TDP43 autoregulation, a phenomenon that involves the nuclear retention and destabilization of *TARDBP* transcripts upon their recognition by TDP43[20,59,219–221]. m6A marks located within the TBR influence TDP43 binding and are crucial for proper autoregulation of the protein. In light of recent data highlighting the possible contribution of TDP43 autoregulation to ALS/FTD pathogenesis[221,238], these observations draw attention to the potential importance of m6A modifications for physiological TDP43 function as well as its dysfunction in disease.

Among different tissue types, the central nervous system displays some of the highest baseline levels of m6A RNA, and these modifications are critical for proper neuronal development and maturation[91,239–241]. Total m6A levels within the nervous system rise with age, and previous evidence suggests that neurodegenerative diseases independent of ALS and FTD are likewise characterized by RNA hypermethylation[196,242]. We observed significant hypermethylation in end-stage tissue from ALS patients compared to age-matched controls; although we do not believe age is responsible for the observed RNA hypermethylation, other factors may be playing a part. Since astrocytosis and microgliosis are common features of ALS as well as other neurodegenerative diseases, it is possible that much of the observed RNA hypermethylation arises not because of a primary disease mechanism, but instead because of secondary neuroinflammation[243].

Importantly, YTHDF2 was also the only m6A-related factor that emerged from our limited CRISPR-based screen for modulators of TDP43-mediated toxicity. This platform, which combines longitudinal fluorescence microscopy, automated survival analysis, and single-cell gene knockout, allowed us to rapidly interrogate most m6A readers, writers, and erasers for their effects on neuronal survival. In doing so, we found that knockout or knockdown of the m6A reader *Ythdf2* mitigated TDP43-related toxicity in rodent primary neurons, while YTHDF2 overexpression itself was lethal, and transfection with writers (METTL3 and METTL14) exacerbated TDP43-dependent neuron loss. Furthermore, *YTHDF2* knockdown in human iNeurons carrying ALS/FTD-associated mutations in *TARDBP* and *C9ORF72* prolonged cellular survival. These data, together with the

apparent accumulation of YTHDF2 in ALS spinal cord, suggest that YTHDF2 may be a novel therapeutic target in ALS and FTD.

## 2.5 Methods

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. HEK293T cells are originally female in origin, are easily transfected, and have been transformed with SV40 T-antigen.

CRISPR/Cas9 integration of HaloTag into *TARDBP* locus in HEK293T cells

Oligos complementary to the target region (Table 2.4) were annealed, digested, and ligated into the *BbsI* site in the pX335 vector (Addgene, #42335) according to the protocol available from Addgene. 1.25 µg of each vector was transfected into HEK293T cells using Lipofectamine 2000 (ThermoFisher, #11668019) together with 2.5 µg of a plasmid encoding the HaloTag open reading frame flanked by 400 bp of sequence homologous to regions immediately upstream and downstream of the *TARDBP* start codon, according to the manufacturer's instructions. Following transfection, cells were split at a low density, allowing transfected cells to establish individual colonies. Cells were screened for nuclear fluorescence after incubation with JF635 dye, as described previously[217,244]. Positive cells were carefully scraped/aspirated using a P200 pipet tip and transferred to a new dish. This process was repeated until 100% of cells displayed nuclear fluorescence after JF635 application, and correct integration of the HaloTag cassette into the TARDBP locus was confirmed by PCR and Sanger sequencing.

HaloTag-TDP43 HEK293T cells were cultured on plates coated with 0.1% gelatin (Sigma, #G2500) in DMEM (GIBCO), 10% FBS, 100 units/mL Penicillin/Streptomycin at 37°C in 5% CO2.

iNeuron differentiation

Day 0: Induced pluripotent stem cells were washed in PBS and incubated in prewarmed accutase (Sigma, #A6964) at 37°C for 8 min. Four volumes of E8 media (ThermoFisher, #A1517001) were added to the plate, and the cells were collected and pelleted at 200xg for 5 min. The media was aspirated, and the pellet was resuspended in 1mL of fresh E8 media. Cells were counted using a hemocytometer, diluted, plated at a density of 20,000 cells/mL in E8 media with ROCK inhibitor and incubated at 37°C overnight. Day 1: Media was changed to N2 media (1x N2 Supplement (Gibco, #17502-048), 1x NEAA Supplement (Gibco, #11140-050), 10 ng/mL BDNF (Peprotech, #450-02), 10 ng/mL NT3 (Peprotech, #450-03), 0.2 µg/mL laminin (Sigma, #L2020), 2 mg/mL doxycycline (Sigma, #D3447) in E8 media). Day 2: Media was changed to transition media ((1x N2 Supplement, 1x NEAA Supplement, 10 ng/mL BDNF, 10 ng/ml NT3, 0.2 µg/mL laminin, 2 mg/mL doxycycline in half E8 media, half DMEM F12 (Gibco, #11320-033)). Day 3: Media was changed into B27 media (1x B27 Supplement (Gibco, #17504-044), 1x Glutamax Supplement (Gibco, #35050-061), 10 ng/mL BDNF, 10 ng/mL NT3, 0.2 µg/mL laminin, and 1x Culture One (Gibco, #A33202-01) in Neurobasal-A (Gibco, #12349-015)). On day 6 cells were transduced with the appropriate virus (prepared by University of Michigan Vector Core, Table S2.2) and cells were sustained in the same

culture medium for the remainder of the experiment. Day 14: Imaging began for survival experiments and iNeurons were monitored over the course of 10 days.

Plasmids

pGW1-GFP, pGW1-TDP43(WT)-GFP, pGW1-mApple[32,222], pGW1-Halo, pGW1-TDP43(WT)-Halo[21,217], and pCAGGS-TDPBR-mCherry[21,216,217] were created as previously described.

To generate pGW1-YTHDF2-Halo, the YTHDF2 ORF was PCR amplified from pcDNA-flag-YTHDF2 (Addgene, #52300). The resulting amplicon was digested with AgeI and XbaI and cloned into cut sites upstream of HaloTag insert in pGW1-Halo.

To create pgW1-YTHDF2-2A-GFP, the YTHDF2 ORF was PCR amplified from pcDNA-flag-YTHDF2 (Addgene, #52300). The resulting amplicon was digested with KpnI and SalI and cloned into cut sites upstream of the 2A in the pGW1-2A-GFP vector.

pCAGGS-TDPBR (mTBR)-mCherry, was created via site directed mutagenesis from pCAGGS-TDPBR-mCherry using the Pfu Ultra high-fidelity polymerase (Agilent Technologies, #600380) according to manufacturer's protocols to change the A into a G in clip34nt sequence in the TDPBR.

HaloTag immunoprecipitation

HEK293T cells were transfected with pGW1-Halo, pGW1-TDP43-Halo, or pGW1-YTHDF2-Halo using Lipofectamine 2000 (ThermoFisher, #11668019) following the manufacturer's protocol. Approximately forty-eight hours post-transfection, cell pellets were harvested and lysed using a syringe in 100 µL lysis buffer (50 mM Tris-HCl, 150 mm

NaCl, 1% Triton-X100, 0.1% NaDeoxycholate) and incubated on ice for 30 min. Following lysis, tubes were centrifuged at 17000xg for 5 min and pellet discarded. 200 g for lysate was added to 25 µL of HaloTrap beads (Chromotek, #ota-10), prewashed with 500 µL of 10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 mM EDTA centrifuged at 2500xg for 5 minutes. Lysate and bead mix was incubated rotating overnight at 4°C. Following incubation, the mixture was centrifuged at 2500xg for 5 min and supernatant discarded. HaloTrap beads were washed 3 times for 5 min at 2500xg in 500 µL of wash buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Triton-X100, 0.1% NaDeoxycholate, 0.05% IGEPAL). To elute RNA, Trizol (ThermoFisher, #15596026) was added directly to the beads and RNA was isolated using phenol-chloroform extraction.

m6A dot blot

For m6A dot blot, isolated RNA was boiled at 95°C for 3 minutes to denature RNA then immediately chilled on ice. RNA was added in 1 µL drops to a BrightStar nylon membrane (ThermoFisher, #AM10102) and crosslinked at 1200 µJ [x100] for 2 min twice. After crosslinking, membranes were stained with methylene blue (0.04% methylene blue in 0.5M sodium acetate pH 5.2) until circles were visible (approximately 5 min). Membranes were washed 3x with water quickly to remove background stain and imaged. Afterwards, membranes were washed several times with water to remove methylene blue staining, then blocked in 5% milk in 1x PBS + 0.02% Tween-20 at room temperature for 1 hour. Primary antibody (rabbit anti-m6A antibody 1:500, Cell Signaling #56593) was added to blocking solution, and membranes rocked overnight at 4°C. Blots were then washed in wash buffer (1x PBS + 0.02% Tween-20) 3x for 5 min and incubated at room

temperature with blocking buffer containing goat anti-rabbit HRP secondary antibody (Jackson Immunoresearch labs, #111-035-003, 1:5000). Blots were rinsed in wash buffer 3x for 5 min, then incubated in Pierce ECL Western blotting substrate (ThermoFisher, #32106) for 1 min before exposure to CL-XPosure™ Film (Fisher, #PI34090). For endogenous Halo-TDP43 pulldowns, the procedure was followed as before starting from cell pellets of Halo-TDP43 HEK293T or unmodified HEK293T cells.

DART-seq

Halo-TDP43 HEK293T cells were transfected with pCMV-APOBEC1-YTH, pCMV-APOBEC1-YTHmut, or pGW1-Halo using Lipofectamine 2000 (ThermoFisher, #11668027) following the manufacturer's protocol. Approximately 48 hours post-transfection, cells were washed with cold PBS (Invitrogen) then crosslinked with UV light (254 nm, 150 mJ/cm$^2$). Following crosslinking, cells were harvested and lysed using a syringe in 100 µL lysis buffer supplemented RNAse inhibitor (Invitrogen, #N8080119) (50 mM Tris-HCl, 150 mM NaCl, 1% Triton-X100, 0.1% NaDeoxycholate) and incubated on ice for 30 min. 200 µg for lysate was added to 25 µL of HaloTrap beads, prewashed with 500 µL of equilibration buffer supplemented with RNAse inhibitor (10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 mM EDTA) centrifuged at 2500xg for 5 minutes. Lysate and bead mix was incubated rotating overnight at 4°C with Baseline Zero DNAse (Fisher Scientific, #NC1424104). Following incubation, the mixture was centrifuged at 2500xg for 5 min and supernatant discarded. HaloTrap beads were washed 3 times for 5 min at 2500xg in 500 µL of wash buffer supplemented with RNAse inhibitor (50 mM Tris-HCl, 150 mM NaCl, 1% Triton-X100, 0.1% NaDeoxycholate, 0.05% IGEPAL). To elute RNA, Trizol was added

directly to the beads and RNA was isolated using phenol-chloroform extraction. Once extracted, RNA was submitted to Advanced Genomics Core at University of Michigan for RNA sequencing.

Next-generation sequencing

cDNA libraries were prepared from Trizol extracted, DNA-digested samples 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 15.6 million unique reads per sample.

DART-seq data processing and analysis

Raw reads were quality and adapter trimmed with Cutadapt (v2.3)[245] with default parameters, and aligned to the GRCh38 human genome assembly with bwa-mem (v0.7.17)[246] with default parameters. Single nucleotide transition analysis was performed as previously described[104]. Briefly, aligned reads were deduplicated, sorted by genome coordinate, parsed to BED format, and collapsed by PCR replicate using the CLIP Tool Kit (CTK) suite[247]. Mutations were extracted and replicates were merged, preserving replicate information. The CTK suite script “CIMS.pl” was then used to generate coverage, mutation rate, and a false discovery statistic. Resulting BED files were filtered in R for entries with a minimum mutation rate of 2, a minimum read count of 10 per replicate, and a mutation/read threshold of 0.1-0.6. Finally, sense C to T (and antisense

G to A) transitions were filtered for presence of the surrounding DRACH motif using custom scripts and the GenomicRanges, BSgenome, Biostrings, and gUtils R packages.

The distribution of mutational transitions was calculated using custom scripts adapted from the MetaPlotR Perl/R suite[248] and visualized with ggplot2. Motif logos were generated before and after DRACH filtration using ggseqlogo. Linear U-G dinucleotide density was calculated using custom R scripts. Briefly, sequences were collapsed to binary representations of UG/GU (1) or non-UG/GU (0) dyads. For UG15 density, a continuous sliding average was calculated in 15-nucleotide windows along the target sequence such that a tract of uniform UG alternation corresponds to a UG15 density of 1. Sliding averages were rescaled to the basepair length of the target sequence and used to generate site-of-interest-centered density plots using ggplot2. Gene ontology analyses were accomplished via STRING or Enrichr[209,249]. Euler diagrams were created using eulerr in R[250].

m6A array

Human spinal cord samples were homogenized in Trizol and RNA was extracted using phenol-chloroform extraction for the m6A mRNA&lncRNA Epitranscriptomic microarray (8x60K, Arraystar, Rockville, MD, USA). 1-3 μg total RNA and m6A spike-in control mixture were added to 300 μL 1xIP buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.1% NP40, 40U/μL RNase Inhibitor) containing 2 μg anti-m6A rabbit polyclonal antibody (Synaptic Systems, #202003). The reaction was incubated with head-over-tail rotation at 4°C for 2 hours. 20 uL DynabeadsTM M-280 Sheep Anti-Rabbit IgG (Invitrogen, #11203D) suspension per sample was blocked with freshly prepared 0.5% BSA at 4°C

for 2 hours, washed three times with 300 µL 1xIP buffer, and resuspended in the total RNA-antibody mixture prepared above. The RNA binding to the m6A-antibody beads was carried out with head-over-tail rotation at 4°C for 2 hours. The beads were then washed three times with 500 µL 1xIP buffer and twice with 500 µL Wash buffer (50 mM Tris-HCl, pH7.4, 50 mM NaCl, 0.1% NP40, 40 U/µL RNase Inhibitor (Enzymatics, #Y9240L)). The enriched RNA was eluted with 200 µL Elution buffer (10 mM Tris-HCl, pH7.4, 1 mM EDTA, 0.05% SDS, 40U Proteinase K) at 50°C for 1 hour. The RNA was extracted by acid phenol-chloroform and ethanol precipitated. The “IP” RNA and “Sup” RNAs were added with equal amount of calibration spike-in control RNA, separately amplified and labeled with Cy3 (for “Sup”) and Cy5 (for “IP”) using Arraystar Super RNA Labeling Kit (Arraystar, #AL-SE-005). The synthesized cRNAs were purified by RNeasy Mini Kit (QIAGEN, #74105). The concentration and specific activity (pmol dye/µg cRNA) were measured with NanoDrop ND-1000. 2.5 µg of Cy3 and Cy5 labeled cRNAs were mixed. The cRNA mixture was fragmented by adding 5 µL 10x Blocking Agent and 1 µL of 25x Fragmentation Buffer, heated at 60°C for 30 min, and combined with 25 µL 2x Hybridization buffer. 50 µL hybridization solution was dispensed into the gasket slide and assembled to the m6A-mRNA & lncRNA Epitranscriptomic Microarray slide. The slides were incubated at 65°C for 17 hours in an Agilent Hybridization Oven. The hybridized arrays were washed, fixed and scanned using an Agilent Scanner G2505C. Comparisons with DART-seq results (Fig. 2.4G) included a limited set of transcripts (2184 out of 5646) that were expressed in both human spinal cord and HEK293T cells [normalized transcripts per million (TPM) > 2], based on datasets made available through the Human Protein Atlas.

Longitudinal microscopy and automated survival analysis

Cortices from embryonic day (E)19-20 Long-Evans rat embryos were dissected and disassociated, and primary neurons were plated at a density of $6x10^5$ cells/ml in 96-well plates, as described previously[21,216,222,251]. For CRISPR candidate screen, on *in vitro* day (DIV) 4, neurons were transfected with 25 ng of plasmids containing sgRNAs for respective genes, 50 ng of pGW1-mApple or pGW1-TDP43-mApple, and 25 ng of pGW1-EGFP to mark cell bodies using Lipofectamine 2000 (Invitrogen #52887. Following transfection, cells were placed in Neurobasal Complete Media (Neurobasal (Gibco, #21103-049), 1x B27, 1x Glutamax, 100 units/mL Pen Strep (Gibco, #15140-122)) and incubated at 37°C in 5% CO2.

m6A CRISPR candidate screen

Oligos complementary to m6A pathway component genes were created using ChopChop[252–254] (https://chopchop.cbu.uib.no/) then annealed, digested, and ligated into the *BbsI* site of pSpCas9(BB)-2A-GFP plasmid (Addgene, #48138) according to manufacturer's protocol. 25 ng of plasmids containing sgRNAs for respective genes were transfected into rat primary neurons on DIV 4 using Lipofectamine 2000 (ThermoFisher, #11668027) along with 50 ng of TDP43(WT)-EGFP or mApple then survival was measured over the course of 10 days. Outcomes were calculated based on Cox proportional hazard values and related to expression of TDP43(WT)-EGFP to determine if knockout of target was beneficial or toxic.

Neurons were imaged as described previously[21,216,222,251] using a Nikon Eclipse Ti inverted microscope with PerfectFocus3a 20X objective lens and either an Andor iXon3 897 EMCCD camera or Andor Zyla4.2 (+) sCMOS camera. A Lambda 421 multi-LED light source (Sutter) with 5 mm liquid light guide (Sutter) was used to illuminate samples, and custom scripts written in Beanshell for use in µManager controlled all stage movements, shutters, and filters. Custom ImageJ/Fiji macros and Python scripts were used to identify neurons and draw both cellular and nuclear regions of interest (ROIs) based upon size, morphology, and fluorescence intensity. Fluorescence intensity of labeled proteins was used to determine protein localization or abundance. Custom Python scripts were used to track ROIs over time, and cell death marked a set of criteria that include rounding of the soma, loss of fluorescence and degeneration of neuronal processes.

Immunocytochemistry

HEK293T cells were plated on glass coverslips (Fisher Scientific, #1254580) and grown overnight at 37°C with 5% $CO_2$. The following day, JF646 (Promega, #GA1120) was added to HEK293T media at 1:10,000 dilution and added to cells for 30 minutes. After 30 minutes, JF646 containing media was removed and washed twice with regular HEK293T media for 15 minutes each. Following wash out, coverslips were washed once with 1x PBS (Gibco, #14200-075) and fixed in 4% PFA for 10 minutes. Cells were then permeabilized with PBS + 0.1% Triton-X-100 for 20 minutes at room temperature, treated with 10 mM glycine for 20 minutes at room temperature, and then blocked in blocking buffer (PBS + 0.1% Triton-X-100, 2% fetal calf serum, and 3% BSA) for 1 hour at room temperature. Coverslips were then incubated overnight with blocking buffer + rabbit anti-

TDP43 antibody (Proteintech, #10782-2-AP) at 1:500 to stain for TDP43. Following primary incubation, coverslips were washed three times with PBS for 5 minutes then stained with secondary antibody goat anti-rabbit AF488 (ThermoFisher, #A-11008) at 1:250 for 1 hour at room temperature in PBS + 0.1% Triton-X-100, 2% fetal calf serum, and 3% BSA. After secondary incubation, coverslips were washed three times for 5 minutes in PBS then stained with Hoechst 1:20,000 to mark nucleus. The coverslips were mounted on coverslips and allowed to dry overnight before imaging. Coverslips were imaged using confocal mode on ONI Nanoimager.

Immunohistochemistry

Immunostaining was accomplished using the Dako Autostainer Link 48 (Agilent, USA). Anti-YTHDF2 antibody (Proteintech, #24744-1-AP, 1:300) or anti-m6A antibody (Synaptic Systems #202003, 1:100) were 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 0.25 microns per 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[255] and mean intensity calculated for stained cells.

RT-PCR and quantitative RT-PCR

Total RNA was extracted using Trizol following the manufacturer's protocol. To synthesize cDNA, 500 ng of total RNA was used in a 20 µL reaction volume with the Bio-Rad iScript cDNA synthesis kit (Bio-Rad, #1708890) according to the manufacturer's protocol. The reactions were incubated at 25°C for 5 min, 46°C for 20 min, and 95°C for 1 min. For quantitative RT-PCR (qRT-PCR), reactions were carried out using Step One Plus Realtime PCR system (Applied Biosystems). Reactions were carried out using PowerUp™ SYBR™ Green Master Mix (ThermoFisher #A25742), with 1 µM primers, and 1 µL cDNA, according to manufacturer's protocol. Relative gene expression was calculated using the ΔΔCt method. Values obtained from qRT-PCR were plotted in GraphPad Prism.

CFTR minigene assay

HEK293 cells were co-transfected with EGFP or TDP43(WT) and the *CFTR* minigene using Lipofectamine 2000. 48h post-transfection, total RNA was extracted using Trizol according to manufacturer's protocol. cDNA was synthesized using the Bio-Rad iScript cDNA synthesis kit according the manufacturer's protocol, and RT-PCR was accomplished as described previously[256].

Purification of recombinant TDP43

TDP43(WT) was expressed in BL21 DE3 E. coli cells from the plasmid pE-6xHis-SUMO-TDP43(WT) (a gift from Dr. James Shorter). Induction was carried out with 1 mM isopropyl-b-D-1-thiogalactopyranoside and cells were grown at 15°C for 16h. Cell pellets were resuspended in lysis buffer (50 mM HEPES, 2% Triton X-100, 300 mM NaCl, 5% glycerol, 50 mM imidazole, 2 mM BME, EDTA-free protease inhibitor cocktail, 5 mM pepstatin, and 20 mg/mL lysozyme) and incubated on ice for 30 minutes. Following sonication on ice, cell lysates were centrifuged for 20 min at 11,000 x g at 4°C. Recombinant protein was purified by binding to Ni-NTA resin (Qiagen #30210), rinsed with 25 mL of wash buffer 4 times (50 mM HEPES, 2% Triton X-100, 300 mM NaCl, 5% glycerol, 50 mM imidazole, and 2 mM BME), and released with 3 mL of elution buffer (50 mM HEPES, 500 mM NaCl, 300 mM imidazole, 5% glycerol, and 5 mM DTT) at room temperature (RT), collecting five 2 mL fractions. Protein was dialyzed twice for 1h in 1 L of final buffer (50 mM HEPES and 500 mM NaCl), and dialyzed once in 1 L of final buffer overnight at 4°C.

Electromobility shift assays (EMSAs)

Binding assays were performed with purified full-length recombinant TDP43 protein and either unmodified or m6A modified ssRNA labeled probes tagged with a 5' 800nm infrared (IR) moiety (IDT; Table 2.4). Binding reactions were performed in binding buffer (12.5 mM HEPES, pH 7.8, 50 mM KCl, 2.5 mM MgCl2, 0.5 mM TCEP, 25 mg/mL BSA, 0.01% NP-40) with 50% glycerol, 1 mg/ml poly-dIdC, 500 pM of labeled probe, and recombinant protein (concentrations indicated in figure legends). Reactions were

incubated on ice for 5 min followed by 25 min at RT. Electrophoresis of 6% acrylamide gels were performed at 60V. Images were acquired using the LI-COR Odyssey platform.

Statistical Analysis

Statistical analysis performed using GraphPad Prism 9 or Superplots[257,258]. For primary neuron survival analysis, the open-source R survival package was used to determine hazard ratios and statistical significance between conditions through Cox proportional hazards analysis[21,217,225,259].

Ethics statement

All vertebrate animal work was approved by the Committee on the Use and Care of Animals (UCUCA) at the University of Michigan. All experiments were performed in accordance with UCUCA guidelines and designed to minimize animal use. Rats (*Rattus norvegicus*) were housed single in chambers equipped with environmental enrichment and cared for by veterinarians from the Unit for Laboratory Animal Medicine at the University of Michigan. All individuals were trained and approved in the care of long-term maintenance of rodents, in accordance with the NIH-supported Guide for the Care and Use of Laboratory Animals. All personnel handling the rats and administering euthanasia were properly trained in accordance with the University of Michigan Policy for Education and Training of Animal Care and Use Personnel. Euthanasia followed the recommendations of the Guidelines on Euthanasia of the American Veterinary Medical Association. Brains from individual pups in each litter were pooled to maximize cell counts prior to plating; as a result, primary cortical neurons used for all studies include an even mix of cells from both male and female pups.

## 2.6 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, Kathy Toy for her expertise with immunohistochemistry, John Moran for assistance with primary neuron survival experiments, and Drs. A. Malik and K. Weskamp for their suggestions.

This work was supported by National Institutes of Health (R01NS097542 and R01NS113943 to SJB; and P30AG072931 to the University of Michigan Brain Bank and Alzheimer's Disease Research Center), 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

S.J.B. and M.M. designed the study. M.M. performed dot blots, immunoprecipitations, qRT-PCR, site directed mutagenesis, transfections, fluorescence

microscopy, primary neuron survival experiments, electromobility shift assays, RNA isolations for DART-seq and the epitranscriptomic array, and immunohistochemical quantifications. N.G. assisted with experimental design, DART-seq and bioinformatics. X.L. assisted with preparation of primary neurons. R.M. wrote the code for image acquisition and neuronal survival analysis. E.M.T. created all knock-in HEK293T cells and iPSCs, in addition to integrating differentiation cassettes for all iPSC lines. M.B. and E.M.T. maintained iPSCs and differentiated iNeurons for survival experiments. M.B. performed and analyzed iNeuron survival studies. S.J.B. and M.M. assembled figures and wrote the manuscript. S.J.B., M.M., E.M.T., M.B., and N.G. edited the manuscript.

## References


1. 1. Charcot J., Joffory A. Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arcives Physiol Neurol Pathol*. 1869;2:744-754.
2. Neumann M. Molecular neuropathology of TDP-43 proteinopathies. *Int J Mol Sci*. 2009;10(1):232-246. doi:10.3390/ijms10010232
3. Ou SH, Wu F, Harrich D, García-Martínez LF, Gaynor RB. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol*. 1995;69(6):3584-3596. http://www.ncbi.nlm.nih.gov/pubmed/7745706.
4. Buratti E, Baralle FE. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front Biosci*. 2008;13:867-878. http://www.ncbi.nlm.nih.gov/pubmed/17981595.
5. Tank EM, Figueroa-Romero C, Hinder LM, et al. Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat Commun*. 2018;9(1):2845. doi:10.1038/s41467-018-05049-z
6. Weskamp K, Barmada SJ. TDP43 and RNA instability in amyotrophic lateral sclerosis. *Brain Res*. 2018;1693(Pt A):67-74. doi:10.1016/j.brainres.2018.01.015
7. Flores BN, Li X, Malik AM, Martinez J, Beg AA, Barmada SJ. An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Rep*. 2019;27(4):1133-1150.e8. doi:10.1016/j.celrep.2019.03.093
8. Klim JR, Williams LA, Limone F, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci*. January 2019:1. doi:10.1038/s41593-018-0300-4
9. Melamed Z, López-Erauskin J, Baughn MW, et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*. 2019;22(2):180-190. doi:10.1038/s41593-018-0293-z
10. Koranda JL, Dore L, Shi H, et al. Mettl14 Is Essential for Epitranscriptomic Regulation of Striatal Function and Learning. *Neuron*. 2018;99(2):283-292.e5. doi:10.1016/J.NEURON.2018.06.007
11. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. 2012;485(7397):201-206. doi:10.1038/nature11112
12. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 2012;149(7):1635-1646. doi:10.1016/j.cell.2012.05.003
13. Sommer S, Lavi U, Darnell JE. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *J Mol Biol*. 1978;124(3):487-499. doi:10.1016/0022-2836(78)90183-3
14. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. 2014;505(7481):117-120. doi:10.1038/nature12730
15. Freibaum BD, Chitta RK, High AA, Taylor JP. Global Analysis of TDP-43 Interacting Proteins Reveals Strong Association with RNA Splicing and

Translation Machinery. *J Proteome Res*. 2010;9(2):1104-1120. doi:10.1021/pr901076y
16. Guo F, Jiao F, Song Z, et al. Regulation of MALAT1 expression by TDP43 controls the migration and invasion of non-small cell lung cancer cells in vitro. *Biochem Biophys Res Commun*. 2015;465(2):293-298. doi:10.1016/J.BBRC.2015.08.027
17. Zhou J, Wan J, Gao X, Zhang X, Jaffrey SR, Qian S-B. Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature*. 2015;526(7574):591-594. doi:10.1038/nature15377
18. Coker H, Wei G, Brockdorff N. m6A modification of non-coding RNA and the control of mammalian gene expression. *Biochim Biophys Acta - Gene Regul Mech*. December 2018. doi:10.1016/J.BBAGRM.2018.12.002
19. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. 2014;505(7481):117-120. doi:10.1038/nature12730
20. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. *N6-Methyladenosine-Dependent RNA Structural Switches Regulate RNA-Protein Interactions*. Vol 518. NIH Public Access; 2015:560. doi:10.1038/NATURE14234
21. Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. *Cell*. 2015;162(6):1299-1308. doi:10.1016/j.cell.2015.08.011
22. Buratti E, Dörk T, Zuccato E, Pagani F, Romano M, Baralle FE. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J*. 2001;20(7):1774-1784. doi:10.1093/EMBOJ/20.7.1774
23. Buratti E, Baralle FE. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator ofCFTR Exon 9 *. *J Biol Chem*. 2001;276(39):36337-36343. doi:10.1074/JBC.M104236200
24. Lukavsky PJ, Daujotyte D, Tollervey JR, et al. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol*. 2013;20(12):1443-1449. doi:10.1038/NSMB.2698
25. Linder B, Grozhik A V, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*. 2015;12(8):767-772. doi:10.1038/nmeth.3453
26. England CG, Luo H, Cai W. HaloTag Technology: A Versatile Platform for Biomedical Applications. *Bioconjug Chem*. 2015;26(6):975-986. doi:10.1021/acs.bioconjchem.5b00191
27. Meyer KD. DART-seq: an antibody-free method for global m6A detection. *Nat Methods*. September 2019:1-6. doi:10.1038/s41592-019-0570-0
28. Hallegger M, Chakrabarti AM, Lee FCY, et al. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell*. 2021;184(18):4680-4696.e22. doi:10.1016/J.CELL.2021.07.018
29. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605. doi:10.1093/NAR/GKAA1074

30. Maor-Nof M, Shipony Z, Lopez-Gonzalez R, et al. p53 is a central regulator driving neurodegeneration caused by C9orf72 poly(PR). *Cell*. 2021;184(3):689-708.e20. doi:10.1016/J.CELL.2020.12.025
31. Ho WY, Navakkode S, Liu F, Soong TW, Ling SC. Deregulated expression of a longevity gene, Klotho, in the C9orf72 deletion mice with impaired synaptic plasticity and adult hippocampal neurogenesis. *Acta Neuropathol Commun*. 2020;8(1). doi:10.1186/S40478-020-01030-4
32. Lambrechts D, Storkebaum E, Morimoto M, et al. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. *Nat Genet 2003 344*. 2003;34(4):383-394. doi:10.1038/ng1211
33. Wang S, Latallo MJ, Zhang Z, et al. Nuclear export and translation of circular repeat-containing intronic RNA in C9ORF72-ALS/FTD. *Nat Commun*. 2021;12(1). doi:10.1038/S41467-021-25082-9
34. Castelli LM, Cutillo L, Souza CDS, et al. SRSF1-dependent inhibition of C9ORF72-repeat RNA nuclear export: genome-wide mechanisms for neuroprotection in amyotrophic lateral sclerosis. *Mol Neurodegener*. 2021;16(1). doi:10.1186/S13024-021-00475-Y
35. Koh CWQ, Goh YT, Goh WSS. Atlas of quantitative single-base-resolution N6-methyl-adenine methylomes. *Nat Commun*. 2019;10(1):5636. doi:10.1038/s41467-019-13561-z
36. Polymenidou M, Lagier-Tourenne C, Hutt KR, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14(4):459-468. doi:10.1038/nn.2779
37. Tollervey JR, Curk T, Rogelj B, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci*. 2011;14(4):452-458. doi:10.1038/nn.2778
38. Bhardwaj A, Myers MP, Buratti E, Baralle FE. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res*. 2013;41(9):5062-5074. doi:10.1093/nar/gkt189
39. Barmada SJ, Ju S, Arjun A, et al. Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proc Natl Acad Sci U S A*. 2015;112(25):7821-7826. doi:10.1073/PNAS.1509744112/-/DCSUPPLEMENTAL
40. Weskamp K, Tank EM, Miguez R, et al. Shortened TDP43 isoforms upregulated by neuronal hyperactivity drive TDP43 pathology in ALS. *J Clin Invest*. 2020;130(3):1139-1155. doi:10.1172/JCI130988
41. Liao S, Sun H, Xu C. YTH Domain: A Family of N6-methyladenosine (m6A) Readers. *Genomics Proteomics Bioinformatics*. 2018;16(2):99-107. doi:10.1016/J.GPB.2018.04.002
42. Ayala YM, De Conti L, Avendaño-Vázquez SE, et al. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J*. 2011;30(2):277-288. doi:10.1038/emboj.2010.310
43. Budini M, Buratti E. TDP-43 autoregulation: Implications for disease. *J Mol Neurosci*. 2011;45(3):473-479. doi:10.1007/S12031-011-9573-8/FIGURES/1
44. Eréndira Avendaño-Vázquez S, Dhir A, Bembich S, Buratti E, Proudfoot N, Baralle FE. Autoregulation of TDP-43 mRNA levels involves interplay between

transcription, splicing, and alternative polyA site selection. *Genes Dev*. 2012;26(15):1679-1684. doi:10.1101/GAD.194829.112
45. Koyama A, Sugai A, Kato T, et al. Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic Acids Res*. 2016;44(12):5820-5836. doi:10.1093/NAR/GKW499
46. Barmada SJ, Serio A, Arjun A, et al. Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat Chem Biol*. 2014;10(8):677. doi:10.1038/NCHEMBIO.1563
47. Archbold HC, Jackson KL, Arora A, et al. TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci Rep*. 2018;8(1):1-18. doi:10.1038/s41598-018-22858-w
48. Barmada SJ, Skibinski G, Korb E, Rao EJ, Wu JY, Finkbeiner S. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci*. 2010;30(2):639-649. doi:10.1523/JNEUROSCI.4988-09.2010
49. Arrasate M, Finkbeiner S. Automated microscope system for determining factors that predict neuronal fate. *Proc Natl Acad Sci U S A*. 2005;102(10):3840-3845. doi:10.1073/PNAS.0409777102
50. Weskamp K, Safren N, Miguez R, Barmada S. Monitoring Neuronal Survival via Longitudinal Fluorescence Microscopy. *J Vis Exp*. 2019;2019(143). doi:10.3791/59036
51. Becker LA, Huang B, Bieri G, et al. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. 2017. doi:10.1038/nature22038
52. Scoles DR, Meera P, Schneider MD, et al. Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. *Nature*. 2017;544(7650):362. doi:10.1038/NATURE22044
53. Scoles DR, Pulst SM. Antisense therapies for movement disorders. *Mov Disord*. 2019;34(8):1112-1119. doi:10.1002/MDS.27782
54. Liu J, Yue Y, Han D, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol*. 2014;10(2):93-95. doi:10.1038/nchembio.1432
55. Sidibé H, Khalfallah Y, Xiao S, et al. TDP-43 stabilizes G3BP1 mRNA: relevance to amyotrophic lateral sclerosis/frontotemporal dementia. *Brain*. 2021;144(11):3461. doi:10.1093/BRAIN/AWAB217
56. Fernandopulle MS, Prestil R, Grunseich C, Wang C, Gan L, Ward ME. Transcription-factor mediated differentiation of human iPSCs into neurons. *Curr Protoc cell Biol*. 2018;79(1):e51. doi:10.1002/CPCB.51
57. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011;72(2):257. doi:10.1016/J.NEURON.2011.09.010
58. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*. 2011;72(2):245-256. doi:10.1016/J.NEURON.2011.09.011
59. Ghasemi M, Brown RH. Genetics of Amyotrophic Lateral Sclerosis. *Cold Spring Harb Perspect Med*. 2018;8(5). doi:10.1101/CSHPERSPECT.A024125

60. Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA*. 1997;3(11):1233. /pmc/articles/PMC1369564/?report=abstract.
61. Ping X-L, Sun B-F, Wang L, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res*. 2014;24(2):177-189. doi:10.1038/cr.2014.3
62. Casafont I, Bengoechea R, Tapia O, Berciano MT, Lafarga M. TDP-43 localizes in mRNA transcription and processing sites in mammalian neurons. *J Struct Biol*. 2009;167(3):235-241. doi:10.1016/J.JSB.2009.06.006
63. Liu N, Parisien M, Dai Q, Zheng G, He C, Pan T. Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA*. 2013;19(12):1848-1856. doi:10.1261/rna.041178.113
64. Schwartz S, Mumbach MR, Jovanovic M, et al. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep*. 2014;8(1):284-296. doi:10.1016/j.celrep.2014.05.048
65. Edupuganti RR, Geiger S, Lindeboom RGH, et al. N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. *Nat Struct Mol Biol*. 2017;24(10):870-878. doi:10.1038/nsmb.3462
66. Ke S, Pandya-Jones A, Saito Y, et al. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev*. 2017;31(10):990-1006. doi:10.1101/GAD.301036.117/-/DC1
67. Louloupi A, Ntini E, Conrad T, Ørom UAV. Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell Rep*. 2018;23(12):3429-3437. doi:10.1016/J.CELREP.2018.05.077
68. Wei G, Almeida M, Pintacuda G, et al. Acute depletion of METTL3 implicates N 6-methyladenosine in alternative intron/exon inclusion in the nascent transcriptome.
69. Sun L, Fazal FM, Li P, et al. RNA structure maps across mammalian cellular compartments. *Nat Struct Mol Biol*. 2019;26(4):322. doi:10.1038/S41594-019-0200-7
70. White MA, Kim E, Duffy A, et al. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat Neurosci*. 2018;21(4):552-563. doi:10.1038/S41593-018-0113-5
71. Cui Q, Shi H, Ye P, et al. m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep*. 2017;18(11):2622. doi:10.1016/J.CELREP.2017.02.059
72. Li F, Yi Y, Miao Y, et al. N6-methyladenosine Modulates Nonsense-mediated mRNA Decay in Human Glioblastoma. *Cancer Res*. September 2019:canres.2868.2018. doi:10.1158/0008-5472.CAN-18-2868
73. Widagdo J, Zhao QY, Kempen MJ, et al. Experience-Dependent Accumulation of N6-Methyladenosine in the Prefrontal Cortex Is Associated with Memory Processes in Mice. *J Neurosci*. 2016;36(25):6771. doi:10.1523/JNEUROSCI.4053-15.2016
74. Shafik AM, Zhang F, Guo Z, et al. N6-methyladenosine dynamics in

neurodevelopment and aging, and its potential role in Alzheimer's disease. *Genome Biol*. 2021;22(1):17. doi:10.1186/s13059-020-02249-z
75. Jiang L, Lin W, Zhang C, et al. Interaction of tau with HNRNPA2B1 and N 6-methyladenosine RNA mediates the progression of tauopathy. *Mol Cell*. 2021;81(20):4209-4227.e12. doi:10.1016/J.MOLCEL.2021.07.038
76. Luo J, Xu T, Sun K. N6-Methyladenosine RNA Modification in Inflammation: Roles, Mechanisms, and Applications. *Front cell Dev Biol*. 2021;9. doi:10.3389/FCELL.2021.670711
77. Grimm JB, Muthusamy AK, Liang Y, et al. A general method to fine-tune fluorophores for live-cell and in vivo imaging. *Nat Methods*. 2017;14(10):987. doi:10.1038/NMETH.4403
78. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17(1):10-12. doi:10.14806/EJ.17.1.200
79. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595. doi:10.1093/BIOINFORMATICS/BTP698
80. Shah A, Qian Y, Weyn-Vanhentenryck SM, Zhang C. CLIP Tool Kit (CTK): a flexible and robust pipeline to analyze CLIP sequencing data. *Bioinformatics*. 2017;33(4):566-567. doi:10.1093/BIOINFORMATICS/BTW653
81. Olarerin-George AO, Jaffrey SR. MetaPlotR: a Perl/R pipeline for plotting metagenes of nucleotide modifications and other transcriptomic sites. *Bioinformatics*. 2017;33(10):1563-1564. doi:10.1093/BIOINFORMATICS/BTX002
82. Xie Z, Bailey A, Kuleshov M V., et al. Gene Set Knowledge Discovery with Enrichr. *Curr Protoc*. 2021;1(3):e90. doi:10.1002/CPZ1.90
83. J L. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. 2020. https://cran.r-project.org/package=eulerr.
84. Archbold HC, Jackson KL, Arora A, et al. TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci Rep*. 2018;8(1):4606. doi:10.1038/s41598-018-22858-w
85. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res*. 2014;42(W1):W401-W407. doi:10.1093/NAR/GKU410
86. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res*. 2016;44(W1):W272-W276. doi:10.1093/NAR/GKW398
87. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*. 2019;47(W1):W171-W174. doi:10.1093/NAR/GKZ365
88. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for digital pathology image analysis. *Sci Reports 2017 71*. 2017;7(1):1-7. doi:10.1038/s41598-017-17204-5
89. Ayala YM, Pagani F, Baralle FE. TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS Lett*. 2006;580(5):1339-1344. doi:10.1016/J.FEBSLET.2006.01.052
90. Lord SJ, Velle KB, Dyche Mullins R, Fritz-Laylin LK. Reproducibility: SuperPlots: Communicating reproducibility and variability in cell biology. *J Cell Biol*.

2020;219(6). doi:10.1083/JCB.202001064
91. Goedhart J. SuperPlotsOfData - A web app for the transparent display and quantitative comparison of continuous data from different conditions. *Mol Biol Cell*. 2021;32(6):470-474. doi:10.1091/MBC.E20-09-0583/ASSET/IMAGES/LARGE/MBC-32-470-G003.JPEG
92. Malik AM, Miguez RA, Li X, Ho YS, Feldman EL, Barmada SJ. Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *Elife*. 2018;7. doi:10.7554/ELIFE.35977
93. Ho R, Workman MJ, Mathkar P, et al. Cross-Comparison of Human iPSC Motor Neuron Models of Familial and Sporadic ALS Reveals Early and Convergent Transcriptomic Disease Signatures. *Cell Syst*. 2021;12(2):159-175.e9. doi:10.1016/J.CELS.2020.10.010

## Figures

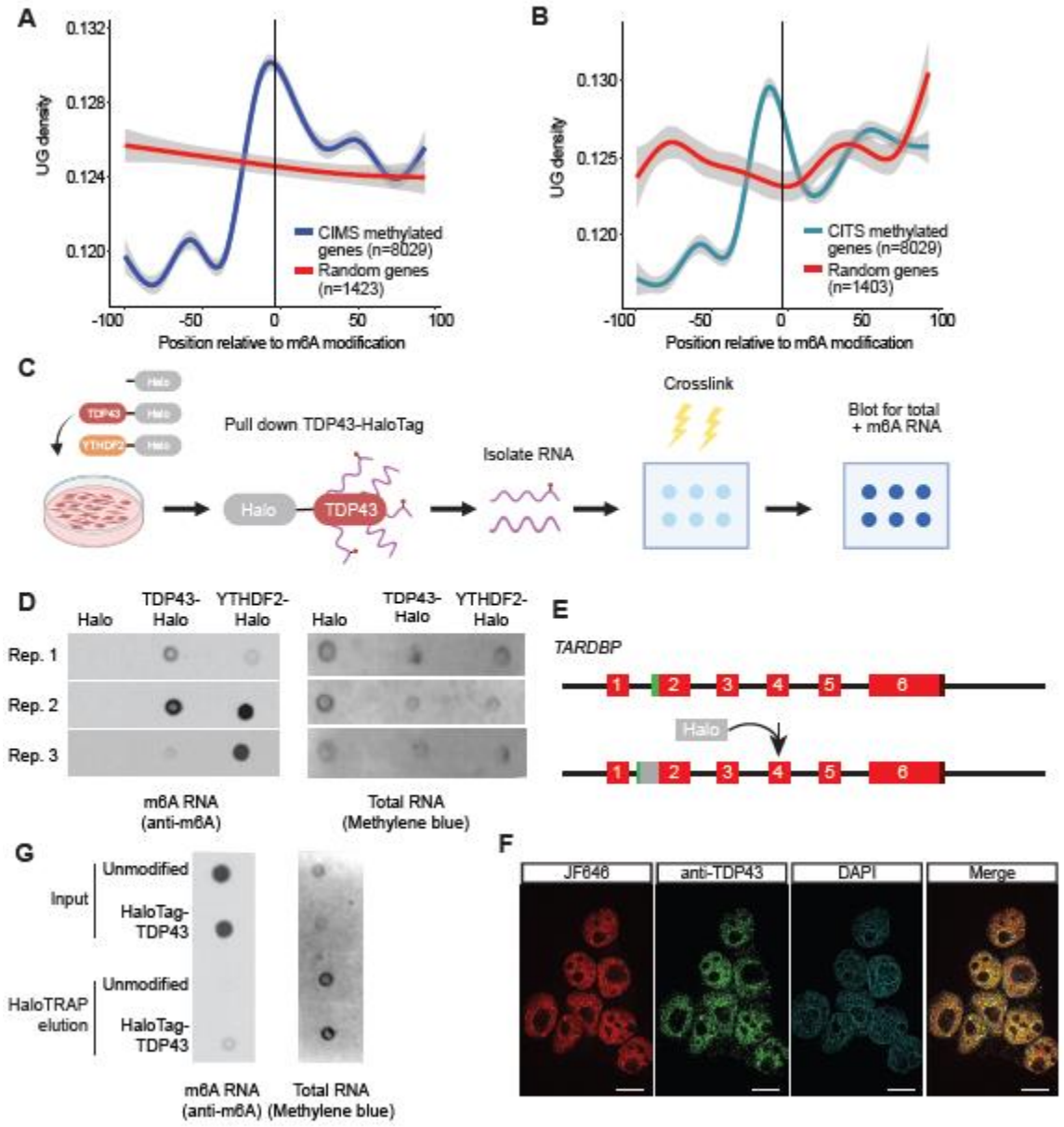



**Figure 2.1: TDP43 binds m6A-modified RNA.**
Density of UG nucleotide sequences 100bp upstream and downstream of m6A modifications identified by cross-linking induced mutation sites (CIMS; **A**) or cross-linking induced truncation sites (CITS; **B**) in relation to random sequences (red line). Grey shading represents 95% confidence regions. (**C**) Schematic of HaloTag immunoprecipitation and dot blot procedure. (**D**) Dot blot for total RNA (detected by methylene blue) or m6A-modified RNA (detected by anti-m6A antibody) isolated by immunoaffinity purification of HaloTag-labeled proteins in HEK293T cells overexpressing HaloTag, TDP43-HaloTag or YTHDF2-HaloTag from 3 biological replicates. (**E**) Diagram illustrating insertion of the HaloTag open reading frame into the endogenous *TARDBP* locus immediately 5' to the TDP43 start codon, resulting in a fusion of HaloTag to the N-terminus of TDP43. (**F**) Halo-TDP43 HEK293T cells labeled live with JF646 Halo dye

(red), then fixed, permeabilized, and immunostained with anti-TDP43 antibody (green) prior to imaging. DAPI (blue) marks the nucleus of each cell. Scale bar = 10µm. (**G**) Dot blot for total RNA (detected by methylene blue) or m6A-modified RNA (detected by anti-m6A antibody) isolated by immunoaffinity purification of endogenous HaloTag-TDP43 or exogenous HaloTag. Additional replicates shown in Sup. Fig. 2.1.

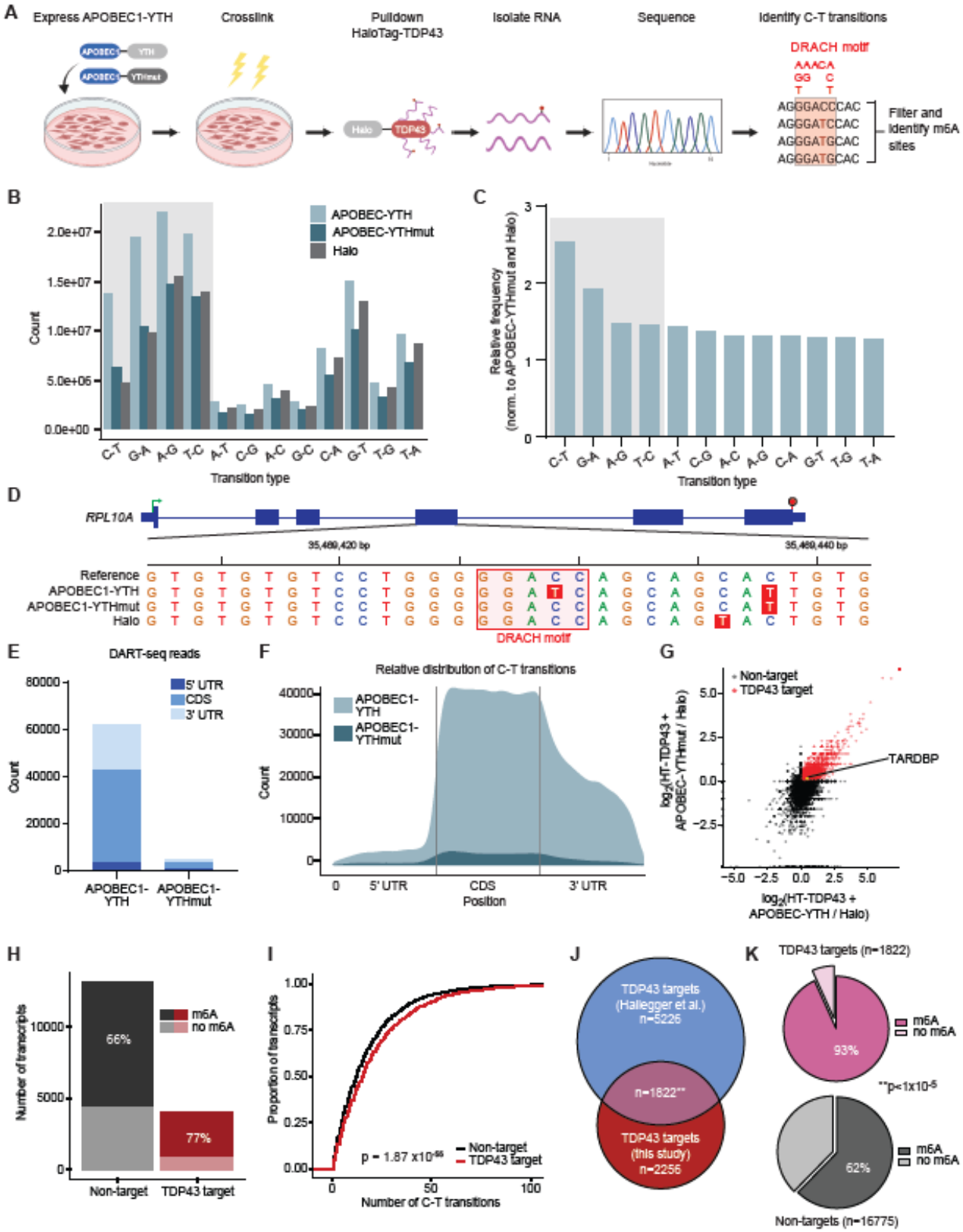



**Figure 2.2: Site-specific identification of m6A-modified TDP43 substrates.**

(**A**) HaloTag-TDP43 immunoprecipitation was followed by DART-seq to delineate m6A sites within TDP43 target RNAs. HaloTag-TDP43 HEK293T cells were transfected with APOBEC1-YTH or APOBEC1-YTHmut and crosslinked before immunoaffinity purification of HaloTag-labeled proteins. Immunoprecipitated RNAs were then sequenced and C-T transitions were identified in the context of DRACH motifs (red shaded box, D=A/G/T, R=A/G, H=A/C/T). Absolute counts (**B**) and relative frequency (**C**) of base pair transitions observed by RNA-seq in each condition. Shaded boxes represent transition types expected from APOBEC1 activity. (**D**) Example m6A sites identified by DART-seq in *RPL10A*. C-T transitions are highlighted in red, and DRACH motifs in pink. Green arrow, transcription start site; red hexagon, transcription stop site; thick blue bars, coding exons; thin blue bars, untranslated region. (**E**) Absolute count and relative distribution (**F**) of DART-seq reads in cells expressing APOBEC1-YTH and APOBEC1-YTHmut. UTR, untranslated region; CDS, coding sequence. (**G**) Scatter plot of TDP43 targets, determined by fold enrichment in precipitated RNA from HaloTag-TDP43 cells (expressing APOBEC1-YTH and APOBEC10YTHmut) compared to cells transfected with HaloTag. Red dots signify transcripts showing $\geq$2-fold enrichment in both APOBEC1-YTH and APOBEC1-YTHmut expressing cells. *TARDBP*, yellow dot, identified as high confidence target. (**H**) Stacked bar graph showing percentage of m6A modified RNA in TDP43 targets (red) and non-targets (black). (**I**) Cumulative distribution of RNA methylation in TDP43 targets (red) and non-targets (black). $p = 1.87x10^{-55}$ by Kolmogorov Smirnov test. (**J**) Euler diagram depicting overlap between TDP43 targets identified in this study, and those identified by TDP43 cross linking and immunoprecipitation followed by RNA-sequencing (CLIP-seq) in HEK293T cells (Hallegger *et al.*, 2021)[28]. **$p=1.5x10^{-117}$, hypergeometric test. (**K**) Pie charts demonstrating the percentage of methylated RNA among TDP43 targets (pink) and non-targets (grey). **$p<1x10^{-5}$ chi-square test.

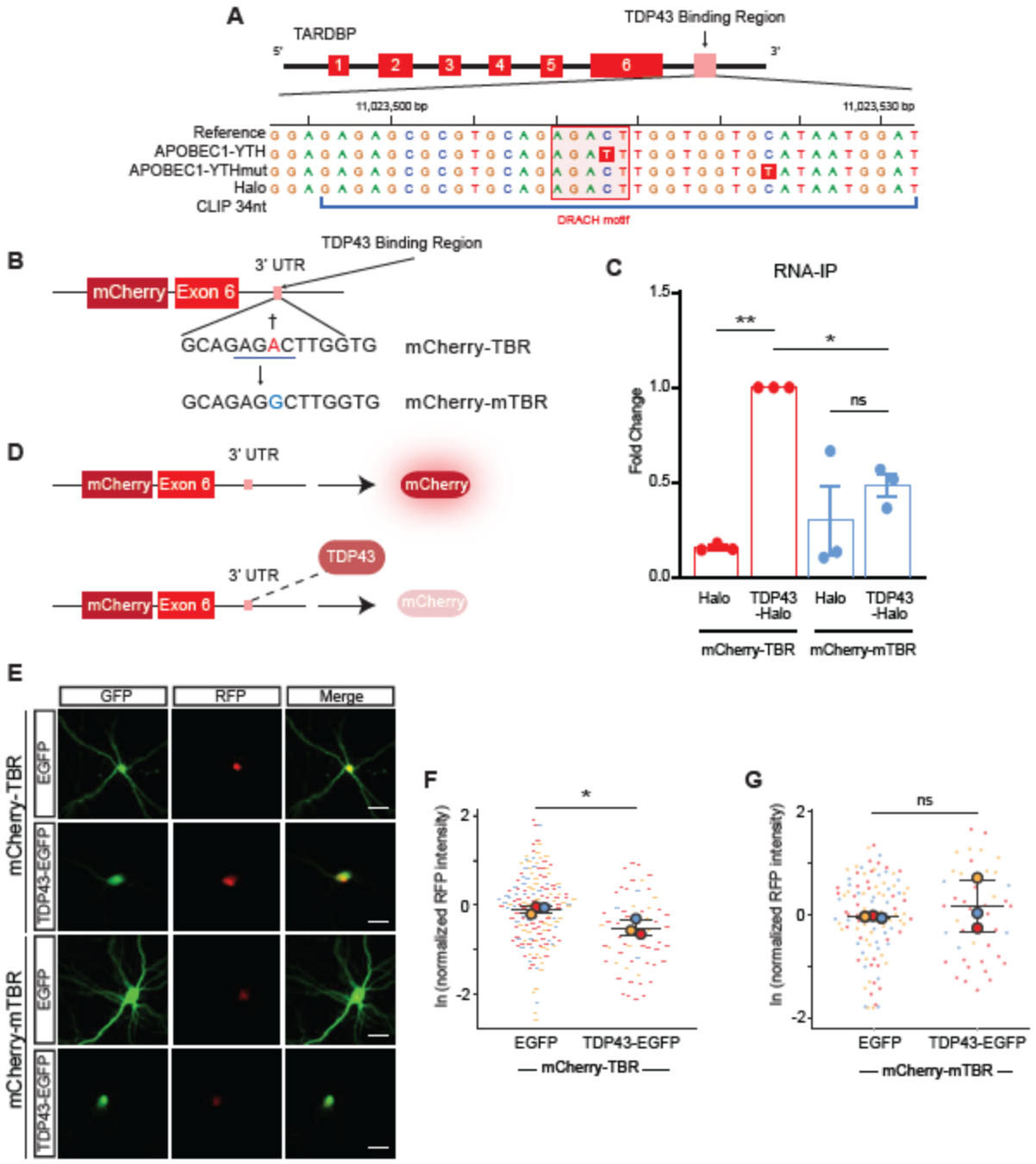



**Figure 2.3: m6A modifications influence TDP43 binding and autoregulation.**
(**A**) *TARDBP* gene map, illustrating TDP43 binding region (TBR), the location of the DRACH motif (pink square), and the C-T transition (red box) identified by DART-seq within this domain, representing an m6A site. (**B**) Schematic of the *TARDBP* minigene reporter, consisting of the mCherry ORF upstream of *TARDBP* exon 6 and 3.4 Kb of the *TARDBP* 3' UTR. The A residue adjacent to the detected C-T transition via DART-seq in the WT reporter (mCherry-TBR) was mutated to a G, precluding methylation the mutant reporter (mCherry-mTBR). Red, methylated residue; blue line, DRACH motif; dagger, C-T transition from DART-seq. (**C**) HaloTag-TDP43 was isolated by immunoaffinity purification from HaloTag-TDP43 HEK293T cells expressing mCherry-TBR or mCherry-

mTBR, and reporter RNA detected in elution fractions by qRT-PCR. (**D**) Outline of TDP43 autoregulation assay. Excess TDP43 binds to the reporter, triggering reporter splicing, destabilization, and reduced mCherry fluorescence. (**E**) Primary rodent neurons were transfected with WT (mCherry-TBR) or mutant (mCherry-mTBR) reporters, together with EGFP or TDP43-EGFP. After 7d, mCherry expression was assessed by fluorescence microscopy. Scale bar= 20 µm. Normalized RFP (mCherry) intensity in primary neurons expressing WT mCherry-TBR reporter (**F**) or mutant mCherry-mTBR (**G**) reporter together with EGFP or TDP43(WT)-EGFP. Cherry-TBR+GFP n= 160, Cherry-TBR+TDP43(WT)-GFP n= 58, Cherry-mTBR+GFP n= 105, Cherry-mTBR+TDP43(WT)-GFP n= 44. Data in **C** plotted as mean ± SD, collected from 3 biological replicates. ns= not significant, *p< 0.05, **p< 0.01; one-way ANOVA with Tukey's test. Data in **F** and **G** plotted as mean ± SD, color coded by biological replicate. ns = not significant, *p < 0.05; Welch's t-test.

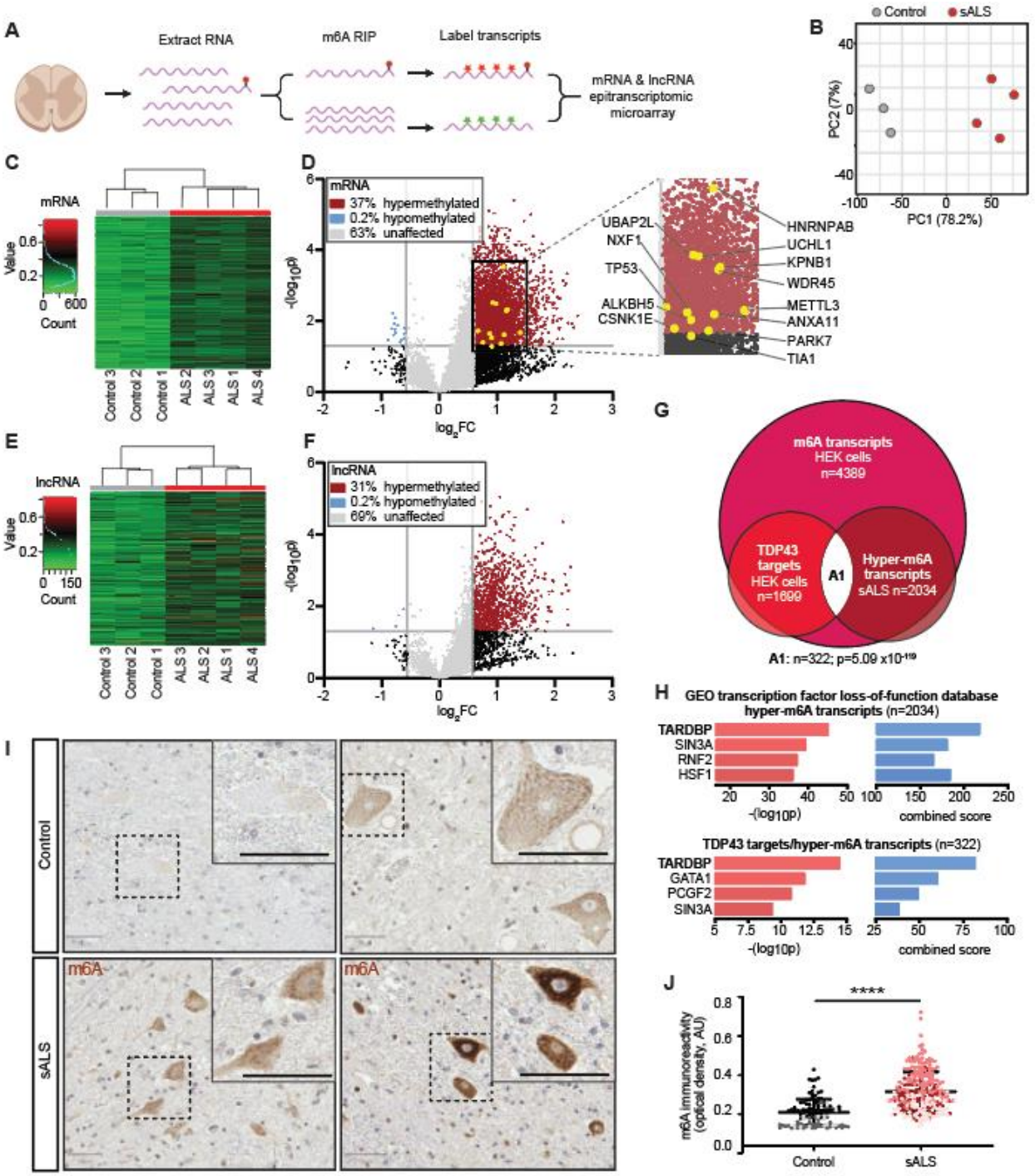



**Figure 2.4: RNA hypermethylation in ALS patient spinal cord.**
(**A**) Genome-wide analysis of RNA methylation via epitranscriptomic array. RNA was extracted from control (n= 3) and sporadic ALS (sALS) patient (n= 4) spinal cord samples, prior to m6A RNA immunoprecipitation. The resulting samples were separated into methylated and non-methylated RNA, then labeled with distinct fluorescent dyes (red and green stars) prior to hybridization, allowing relative quantification of methylation at each annotated locus. (**B**) Principal component analysis (PCA) plot comparing methylation levels of control (grey) and ALS (red) patient samples. (**C**) Hierarchical

clustering of mRNA methylation profiles from control and ALS mRNA samples. (**D**) Volcano plot depicting fold change in mRNA methylation levels in ALS compared to control spinal cord. (**E**) Hierarchical clustering of lncRNA methylation profiles from control ALS lncRNA samples. (**F**) Volcano plot showing fold change in lncRNA methylation levels in ALS compared to control spinal cord. In **D** and **F**, grey horizontal vertical lines represent p= 0.05 and fold change (FC)= 2. (**G**) Euler diagram demonstrating overlap (n= 322, p= $5.09x10^{-119}$, hypergeometric test) among TDP43 substrates and methylated transcripts identified in HEK293T cells, in additional to hypermethylated transcripts determined via m6A array in sALS spinal cord. Comparisons were limited to the subset of transcripts expressed in both HEK293T cells and human spinal cord (nTPM>2). (**H**) Based on comparisons with the GEO transcription factor loss-of-function database via Enrichr[83], there was strong enrichment for TDP43-regulated genes not only among the set of 2034 transcripts hypermethylated in sALS spinal cord, but also among the 322 TDP43 targets that were also hypermethylated in sALS (A1 in **G**). Combined score = ($\log_{10}$p * Z-score). (**I**) Immunohistochemical staining for m6A in control and sALS spinal cord sections. Scale bars= 50 µm. (**J**) Quantification of m6A antibody reactivity in spinal cord neurons from control (n= 110 neurons) and sALS (n= 277 neurons) sections. Plot shows mean +/- SD, color coded by patient. ****p< 0.0001 via Mann-Whitney test.

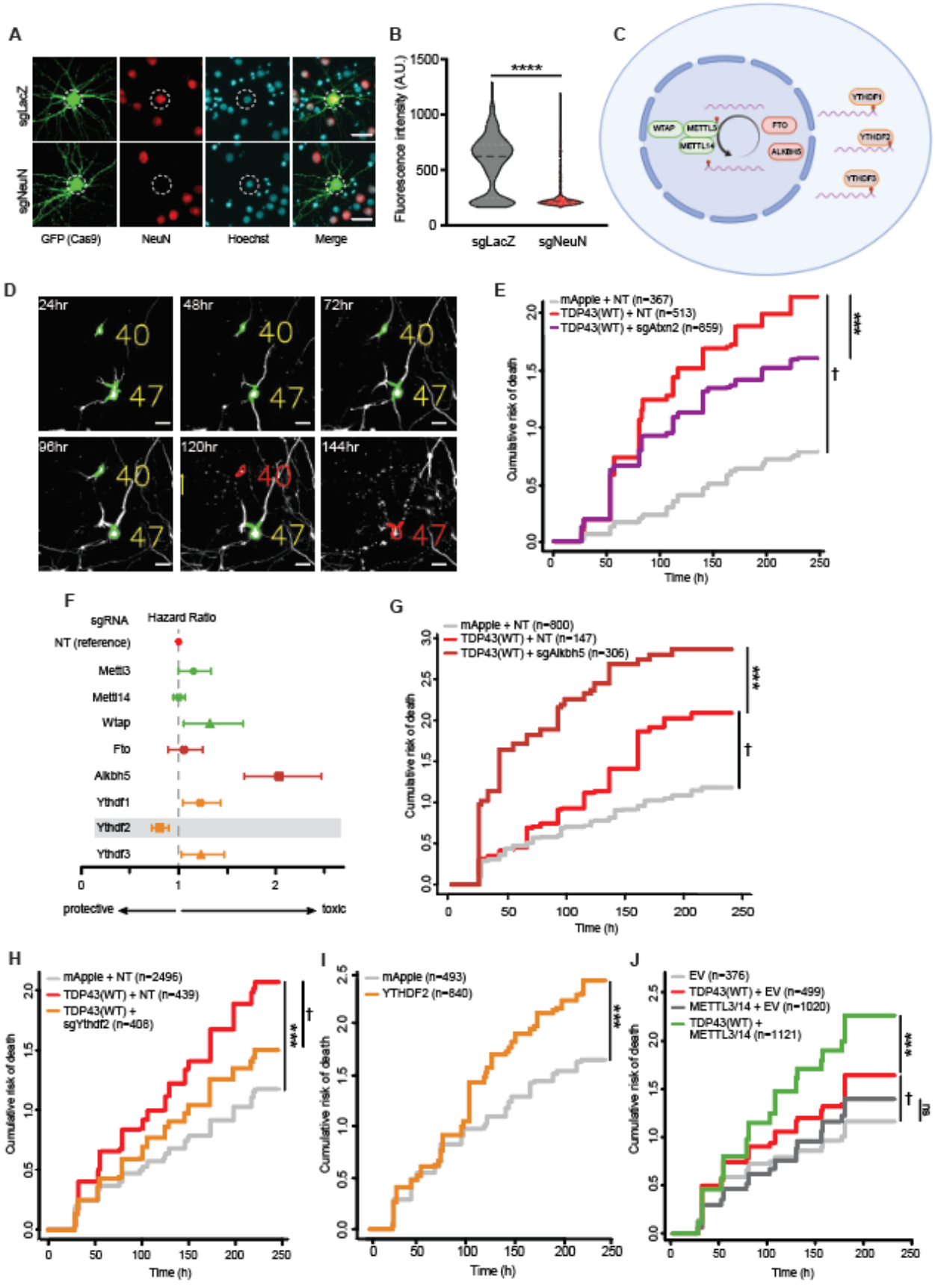

A
sgLacZ
sgNeuN
GFP (Cas9)
NeuN
Hoechst
Merge
B
Fluorescence intensity (A.U.)
1500
1000
500
0
****
sgLacZ
sgNeuN
C
WTAP
METTL3
METTL14
FTO
ALKBH5
YTHDF1
YTHDF2
YTHDF3
D
24hr
48hr
72hr
96hr
120hr
144hr
40
47
E
mApple + NT (n=367)
TDP43(WT) + NT (n=513)
TDP43(WT) + sgAtxn2 (n=659)
Cumulative risk of death
Time (h)
***
†
F
sgRNA
Hazard Ratio
NT (reference)
Metti3
Metti14
Wtap
Fto
Alkbh5
Ythdf1
Ythdf2
Ythdf3
protective
toxic
G
mApple + NT (n=800)
TDP43(WT) + NT (n=147)
TDP43(WT) + sgAlkbh5 (n=306)
Cumulative risk of death
Time (h)
***
†
H
mApple + NT (n=2496)
TDP43(WT) + NT (n=439)
TDP43(WT) + sgYthdf2 (n=408)
Cumulative risk of death
Time (h)
***
†
I
mApple (n=493)
YTHDF2 (n=840)
Cumulative risk of death
Time (h)
***
J
EV (n=376)
TDP43(WT) + EV (n=499)
METTL3/14 + EV (n=1020)
TDP43(WT) + METTL3/14 (n=1121)
Cumulative risk of death
Time (h)
***
†
ns

**Figure 2.5: m6A factors modulate TDP43-dependent neurotoxicity.**
(**A**) Representative images of rodent primary neurons transfected with plasmids expressing Cas9-2A-EGFP and sgRNA targeting the neuronal protein NeuN or negative control (LacZ). 5d after transfection, neurons were fixed and immunostained for NeuN (red). White dashed circles indicate nucleus stained with Hoechst (blue). (**B**) NeuN antibody reactivity measured in EGFP-positive neurons expressing sgLacZ (n= 565) or sgNeuN (n= 654), ****p < 0.0001 by Mann-Whitney. (**C**) Schematic depicting m6A writers (green), erasers (red), and readers (orange) targeted by CRISPR/Cas9. (**D**) Primary neurons expressing EGFP and TDP43-mApple were assessed at regular 24h intervals by fluorescence microscopy, and their survival assessed by automated image analysis. Individual neurons are assigned unique identifiers (yellow number) and tracked until their time of death (red), indicated by cellular dissolution, blebbing, or neurite retraction. Scale bar= 20µm. (**E**) Cumulative hazard plot depicting risk of death for neurons expressing TDP43(WT) + non-targeting (NT) (red line), mApple + NT (grey line), or TDP43(WT) + Atxn2 sgRNA (purple line). †p<2.0 x10-16, Hazard ratio (HR)= 3.45; ***p= 5.81 x10-4, HR= 0.80). (**F**) Forest plot showing HR for TDP43-overexpressing neurons upon knockdown of m6A writers (green), erasers (dark red), and readers (orange), in comparison to nontargeting (NT) control. Dashed line indicates HR= 1, representing the survival of the reference condition, neurons expressing TDP43-mApple and NT sgRNA. Values >1 indicate increased toxicity, whereas values <1 denote relative protection. Error bars represent 95% CI. (**G**) Alkbh5 knockout significantly increases TDP43 associated toxicity. †p=3.11 x10-5, HR= 1.59; ***p= 2.65x10-11, HR= 2.03. (**H**) Ythdf2 knockout significantly extends survival in TDP43-expressing neurons. ***p <2.0 x10-16, HR= 1.69; †p= 6.2 x10-6, HR= 0.71. (I) YTHDF2 overexpression is toxic to neurons. ***p= 3.07x10-5, HR= 1.30. (**J**) METTL3/14 overexpression enhances TDP43-dependent toxicity in neurons. †p = 5.53 x10-4, HR= 1.32; ***p =4.16 x10-6, HR= 1.31. p values in **E, G-J** determined via Cox proportional hazards analysis, with a minimum 3 of biological replicates.

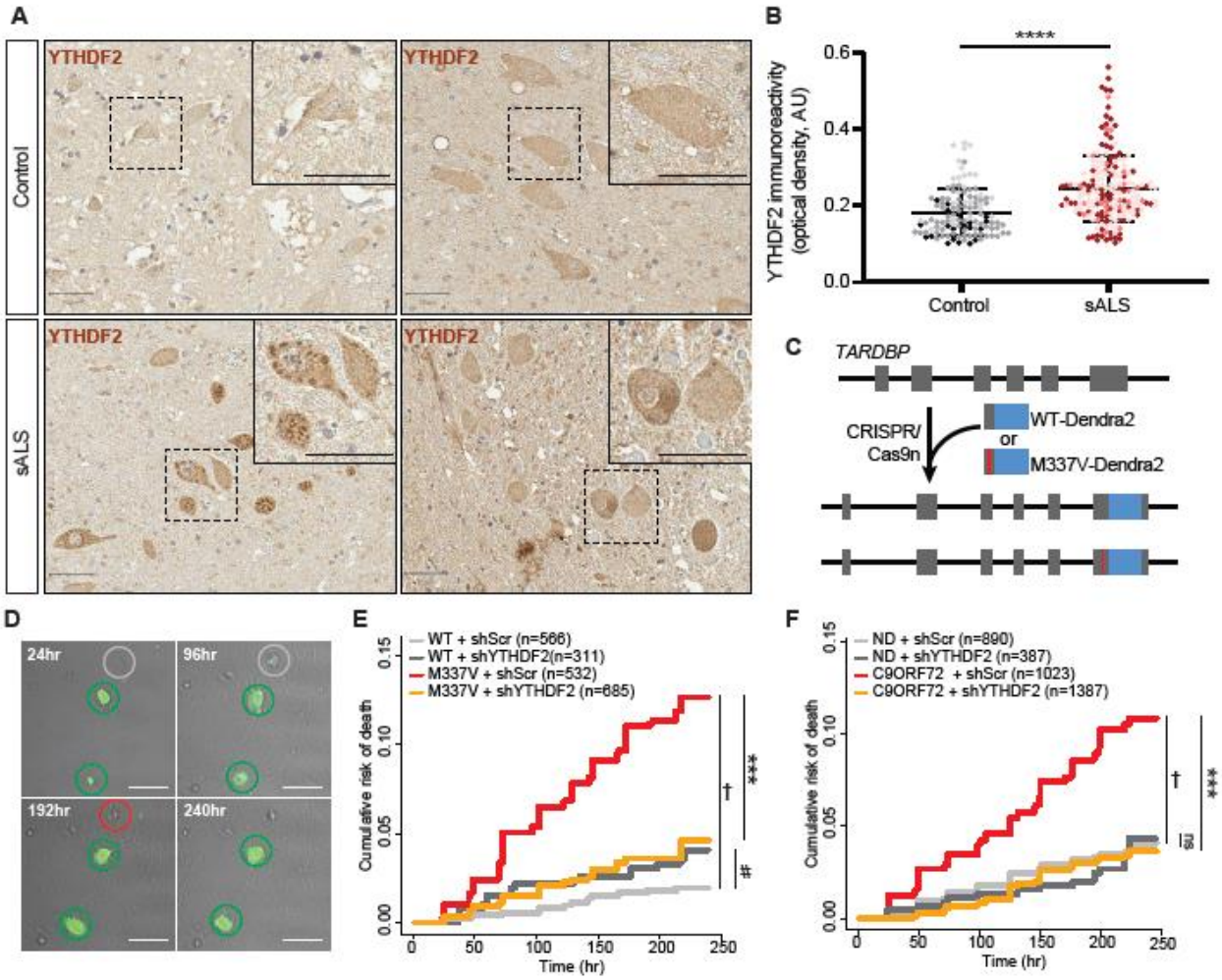
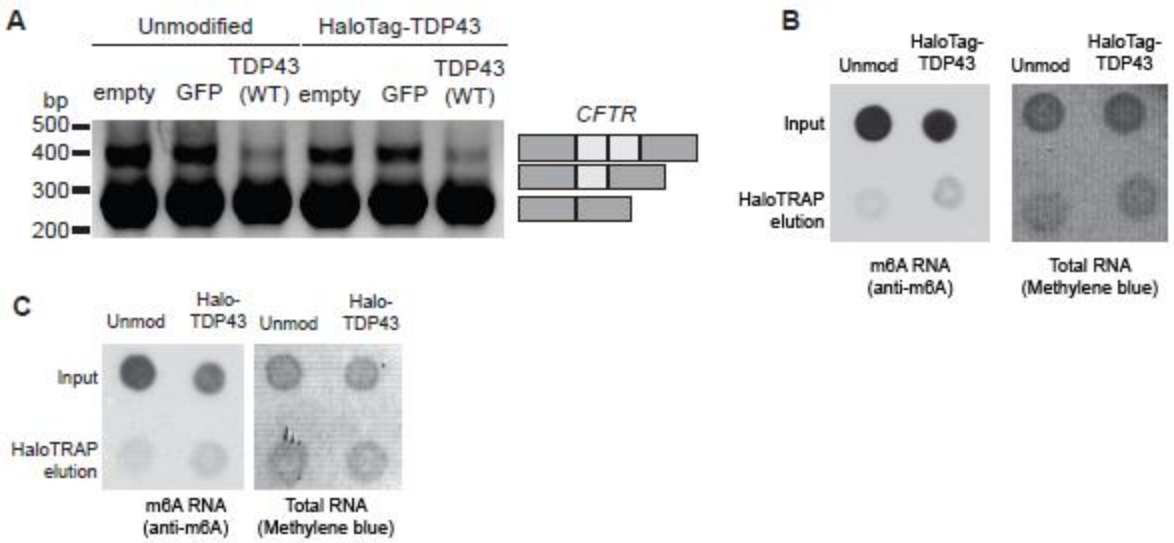
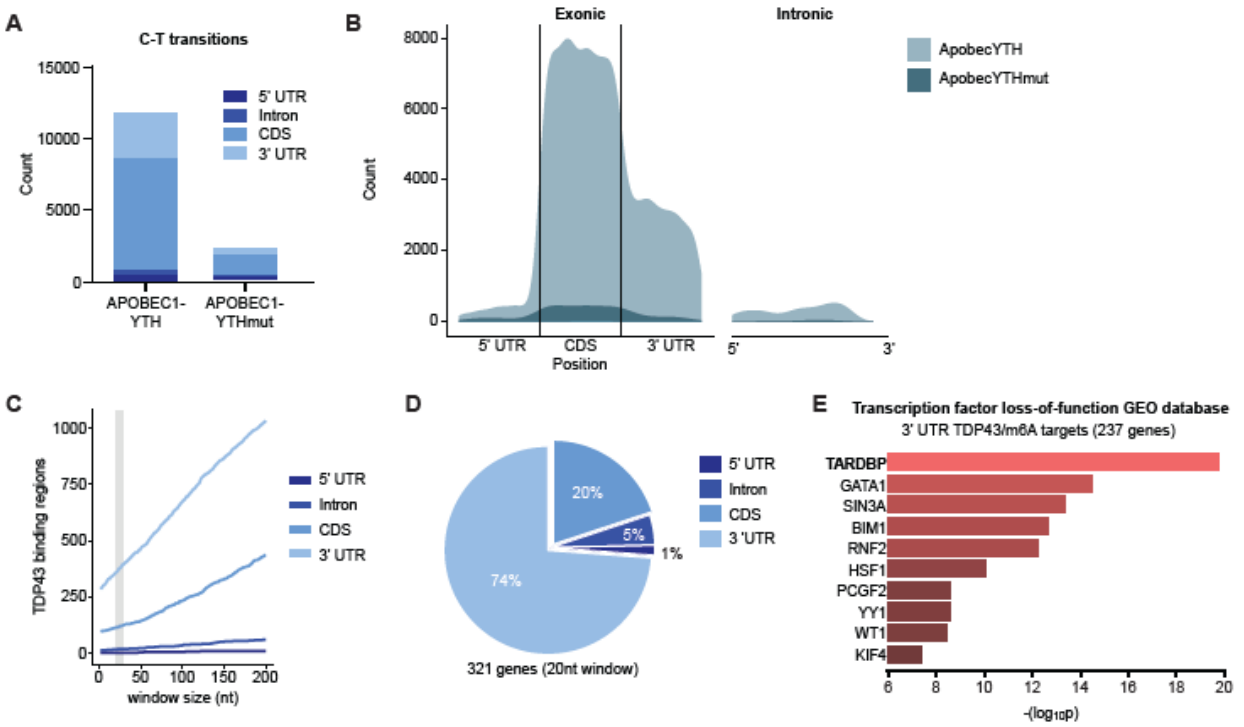



**Figure 2.6: YTHDF2 extends neuronal survival in human neuron disease models.** **(A)** Immunostaining of YTHDF2 in control and sALS patient spinal cord samples. Scale bar= 50 µm. (**B**) Quantification of YTHDF2 immunoreactivity in spinal cord neurons from control (n= 117 neurons) and sALS (n= 193 neurons) samples. Plot shows mean +/- SD, color coded by sample. ****p< 0.0001 via Mann-Whitney test. (**C**) Strategy used to create isogenic iPSCs expressing native TDP43(WT)-Dendra2 or TDP43(M337V)-Dendra2. (**D**) Representative images of untransduced (grey) and transduced (green) iNeurons expressing shRNA against YTHDF2 (shYTHDF2) and a GFP reporter. Time of death (red circles) for each cell is used to determine cumulative risk of death, plotted in (**E**) and (**F**). Scale bar= 20µm. shRNA-mediated knockdown of YTHDF2 significantly extended the survival of TDP43(M337V)-Dendra2 iNeurons (**E**; †p= 8.42x10-12, HR= 6.25; ***p= 4.82x10-9, HR=0.32; #p= 0.08, HR= 1.84) as well as mutant C9ORF72 iNeurons (**F**, †p= 1.42x10-11, HR= 2.85; ***p= 1.42x10-16, HR= 0.32). ns= not significant. Values in (**E, F**) calculated by Cox proportional hazards analysis, with a minimum 3 biological replicates.



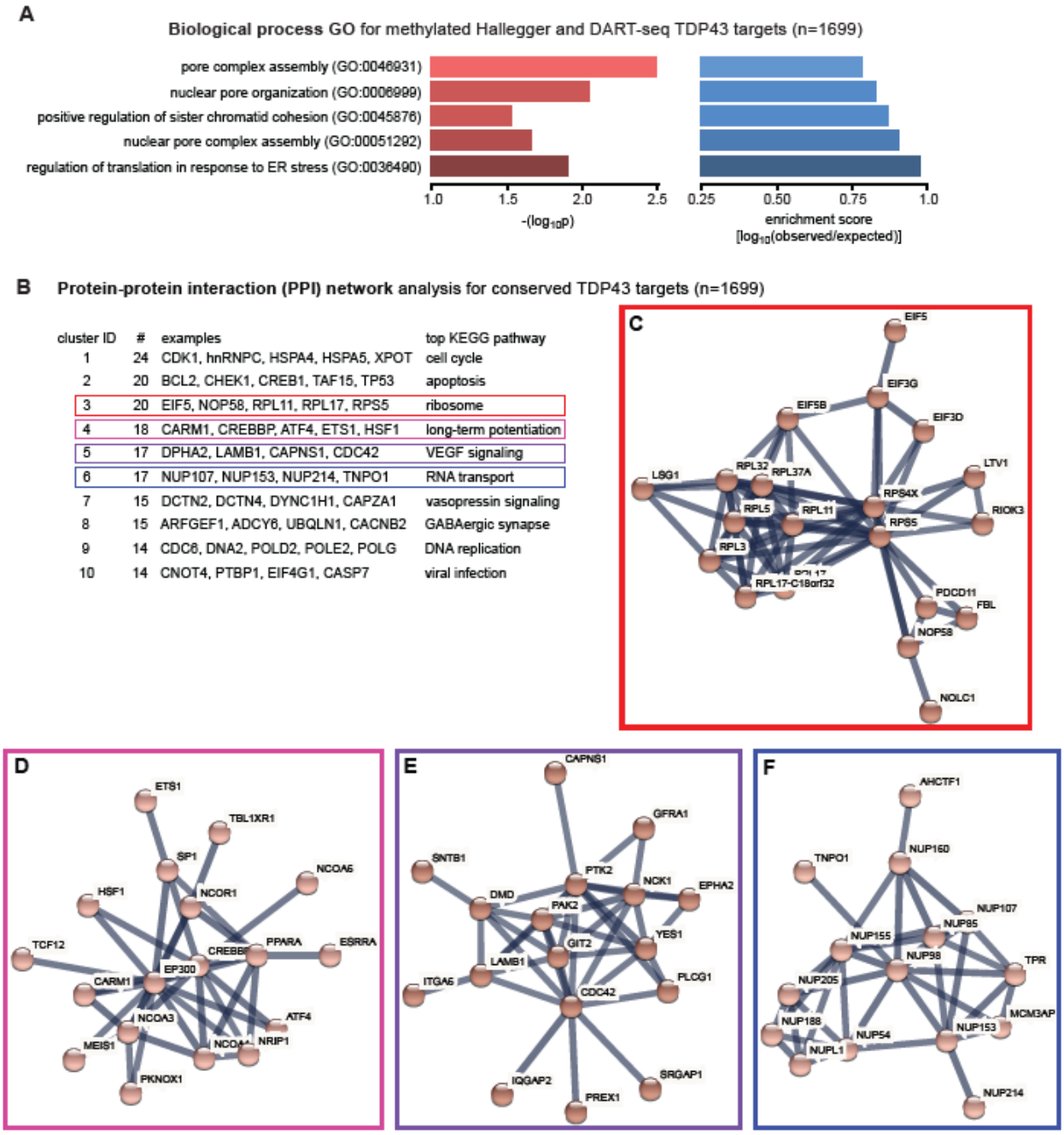

**Supplemental Figure 2.1: HaloTag insertion in TARDBP locus does not affect TDP43 function.**

(**A**) *CFTR* minigene splicing assay[7,33,75] in which unmodified or HaloTag-TDP43 HEK293T cells were transfected with the CFTR minigene along with EGFP or TDP43-EGFP then analyzed by PCR amplification to measure functional TDP43. Correct TDP43 mediated splicing of the reporter results in exon 9 exclusion. (**B**) Additional dot blots for total RNA (detected by methylene blue) or m6A-modified RNA (detected by anti-m6A antibody) isolated by immunoaffinity purification of endogenous HaloTag-TDP43 or exogenous HaloTag using Synaptic Systems anti-m6a (#202003) or (**C**) Millipore Sigma anti-m6A (ABE572-I-100UG) antibodies.



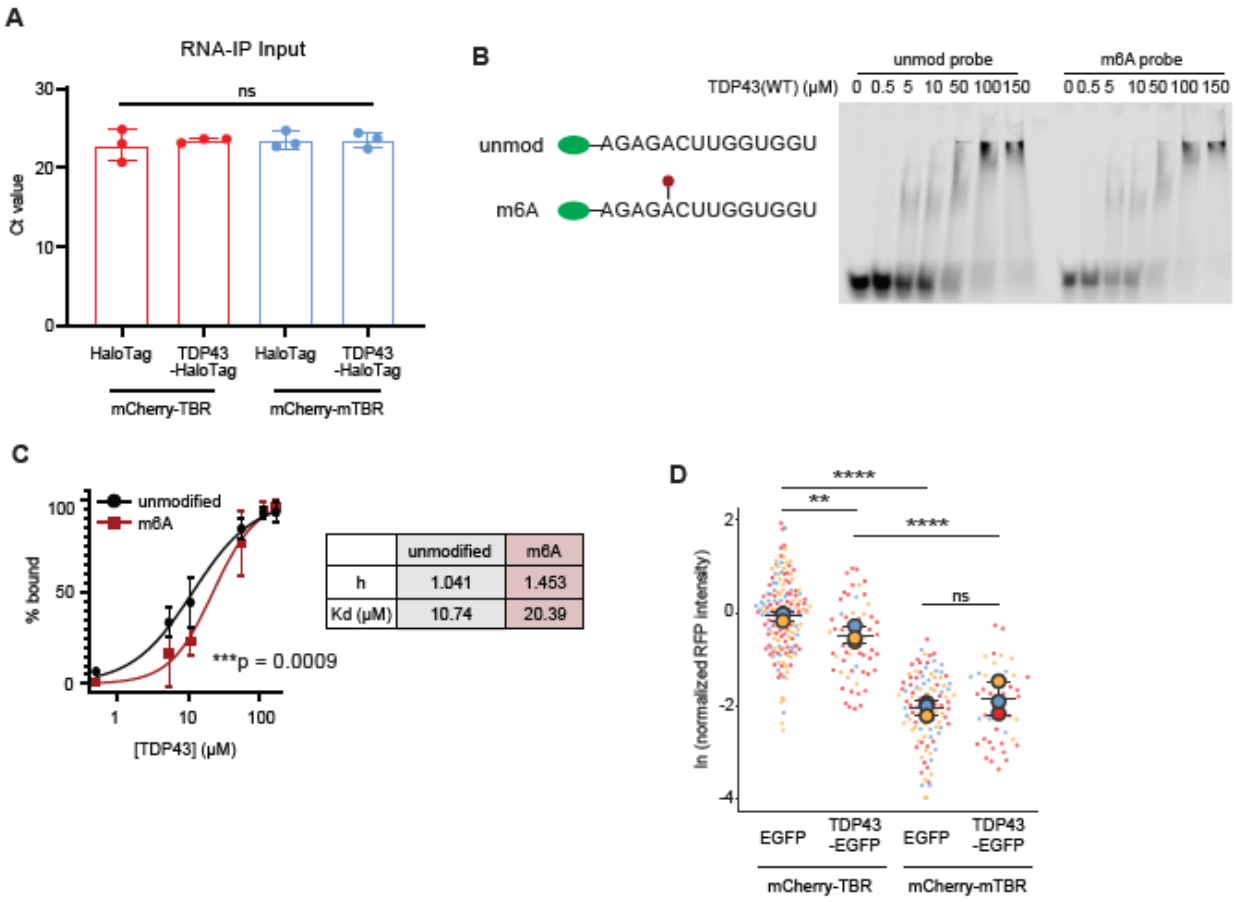

**Supplemental Figure 2.2: Intronic regions of methylated TDP43 targets are devoid of m6A sites.**

Absolute count (**A**) and relative distribution (**B**) of C-T transitions indicative of m6A modifications located within the coding sequence (CDS), 5' untranslated region (UTR), 3' UTR, and introns in cells expressing APOBEC1-YTH and APOBEC1-YTHmut. **(C)** Iterative topological mapping plot depicting the likelihood of finding a TDP43 binding site (as determined by CLIP-seq, Hallegger et al. 2021[208]) within the vicinity of a C-T transition (identified by DART-seq, this study), located at position 0. The magnitude and direction of the slopes for lines corresponding to each gene region indicates a positive relationship between TDP43 binding sites and m6A sites primarily within the 3' UTR, but also within the CDS. (**D**) In the subset of genes exhibiting a TDP43 binding site within 20nt of an m6A site (n=321; grey shading in **C**), this association was most often detected within the 3' UTR (237/321 genes, or 74%). (**E**) The 237 genes showing TDP43 binding sites and m6A sites within the 3'UTR were strongly enriched for genes whose expression is regulated by TDP43, as determined by comparison with the transcription factor loss-of-function GEO database.

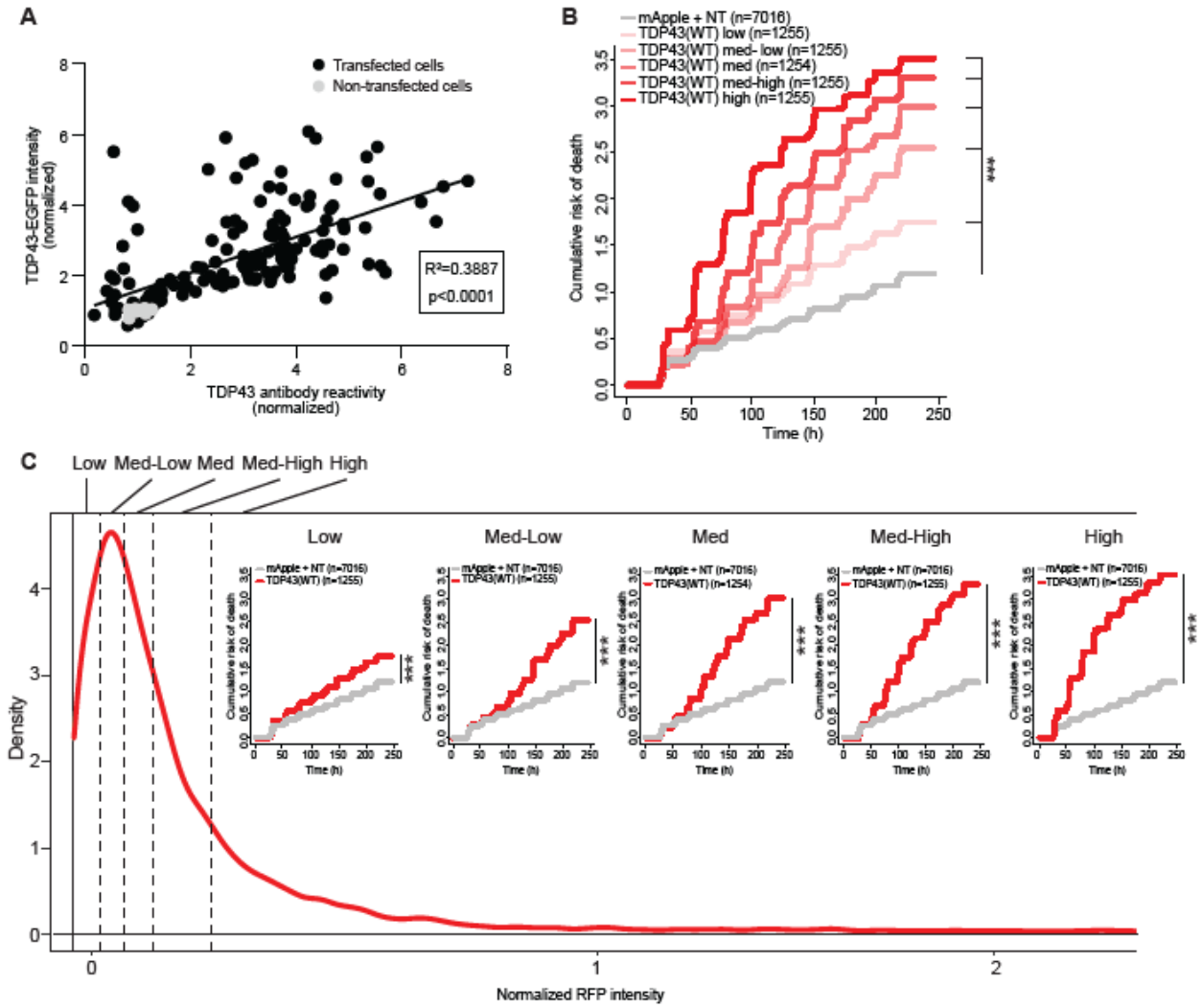



**Supplemental Figure 2.3: Methylated TDP43 targets are enriched in RNA binding and homeostasis.**

(**A**) Gene ontology for biological processes enriched in TDP43 substrate RNAs (n=1699) common to those identified in this study and Hallegger *et al.* 2021[208]. (B) Protein-protein interaction network analysis[209] for shared TDP43 targets shows enrichment of targets associated with ribosome (**C**), long term potentiation (**D**), VEGF signaling (**E**), and RNA transport (**F**).

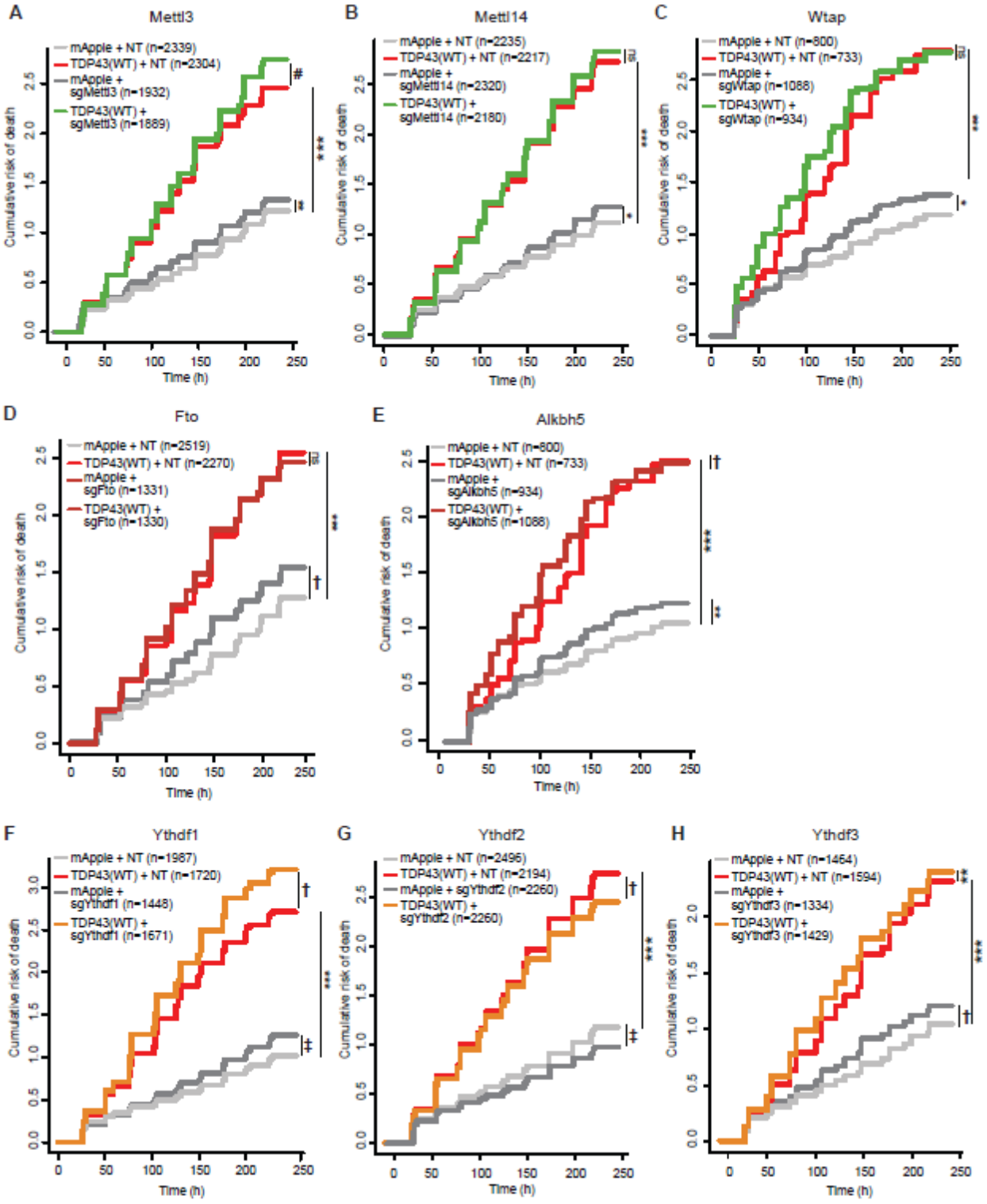



**Supplemental Figure 2.4: m6A modifications alter TDP43 binding capabilities.** (**A**) qRT-PCR cycle threshold (Ct) values from HaloTag-TDP43 HEK293T cells expressing mCherry-TBR or mCherry-mTBR. (**B**) Normalized RFP (mCherry) intensity in primary neurons expressing WT (mCherry-TBR) or mutant (mCherry-mTBR) reporters together with EGFP or TDP43-EGFP. mCherry-TBR+GFP n= 160, mCherry-TBR+TDP43(WT)-GFP n= 58, mCherry-mTBR+GFP n= 105, mCherry-mTBR+TDP43(WT)-GFP n= 44. (**C**) Electromobility shift assay (EMSA) demonstrating binding of recombinant TDP43(WT) to 14nt probe modeled after the TDPBR, with and without m6A modification. Probe concentration was kept constant at 500 pM. (**D**) Percent bound m6A modified (red) or unmodified (black) RNA probe. Data in **A** and **B** plotted as mean ± SD, color coded by biological replicate. ns = not significant; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; one-way ANOVA with Tukey's test. Data in (**C**) plotted as nonlinear regression with Hill slope. ***p < 0.0009; extra-sum-of-squares F test.




**Supplemental Figure 2.5: TDP43-mApple expression is proportional to intensity and toxicity.**

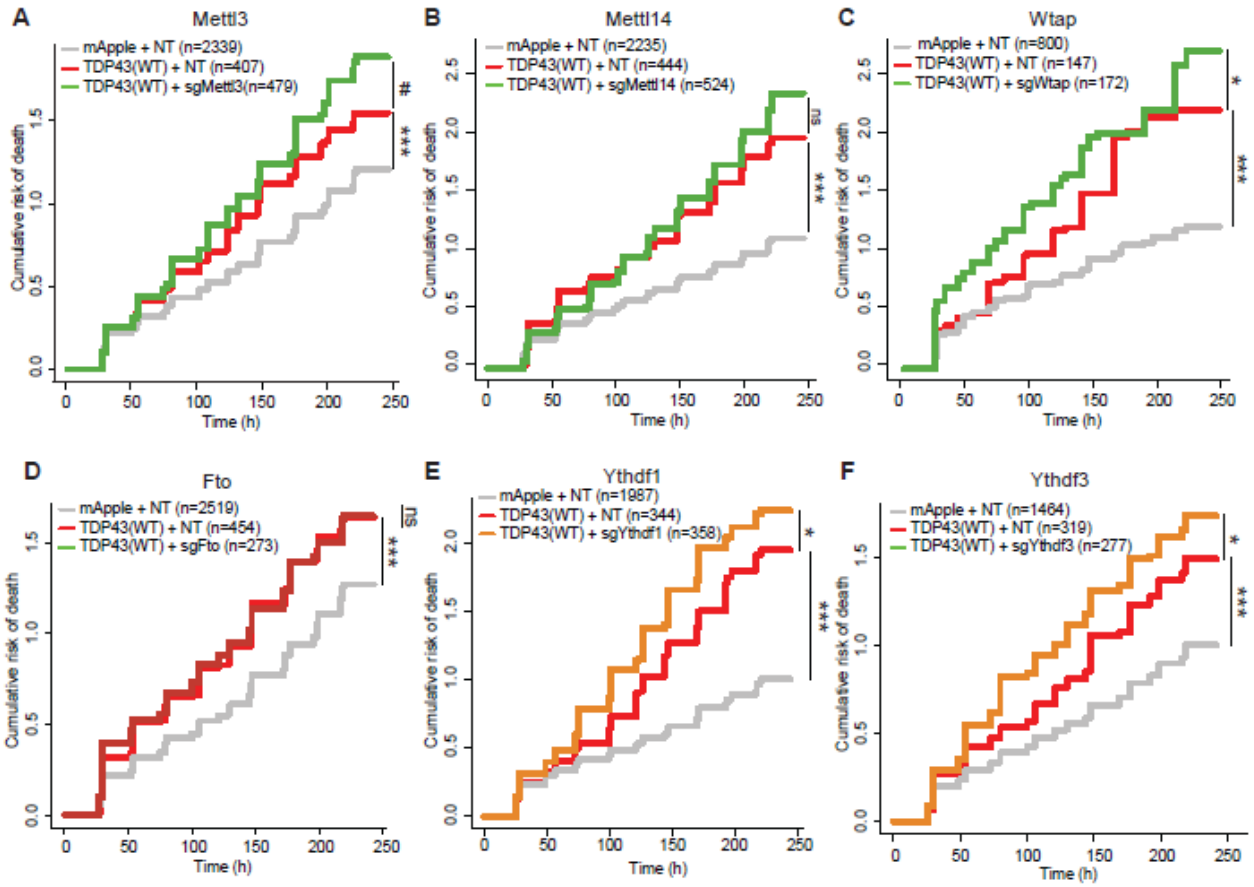
(**A**) Correlation of single-cell TDP43 protein levels, determined by immunostaining, and RFP fluorescence in neurons overexpressing TDP43(WT)-mApple. Black: transfected cells, n= 144; grey: non-transfected cells, n= 20. (**B**) Neurons were stratified into 5 quintiles of equal cell number based on TDP43(WT)-mApple intensity, and their risk of death compared via Cox proportional hazards analysis. n= 7016 for mApple + NT; n= 1254 or 1255 cells per quintile. TDP43(WT)-mApple expression is significantly more toxic at all quintiles compared to mApple control: Low HR= 1.38, Med-Low HR= 1.85, Med HR= 2.18, Med-High HR= 2.47, High HR= 2.89. ***p <2.0 $x10^{-16}$ for all quintiles, Cox proportional hazards analysis. (**C**) Density plot of normalized RFP fluorescence illustrating separation of TDP43(WT)-mApple expression quintiles. Inset graphs correspond to neuronal survival in each quintile, same data as displayed in **B**. ***p <2.0 $x10^{-16}$ for all quintiles, Cox proportional hazards analysis.




**Supplemental Figure 2.6: Knockout of m6A pathway components modulates TDP43 toxicity.**

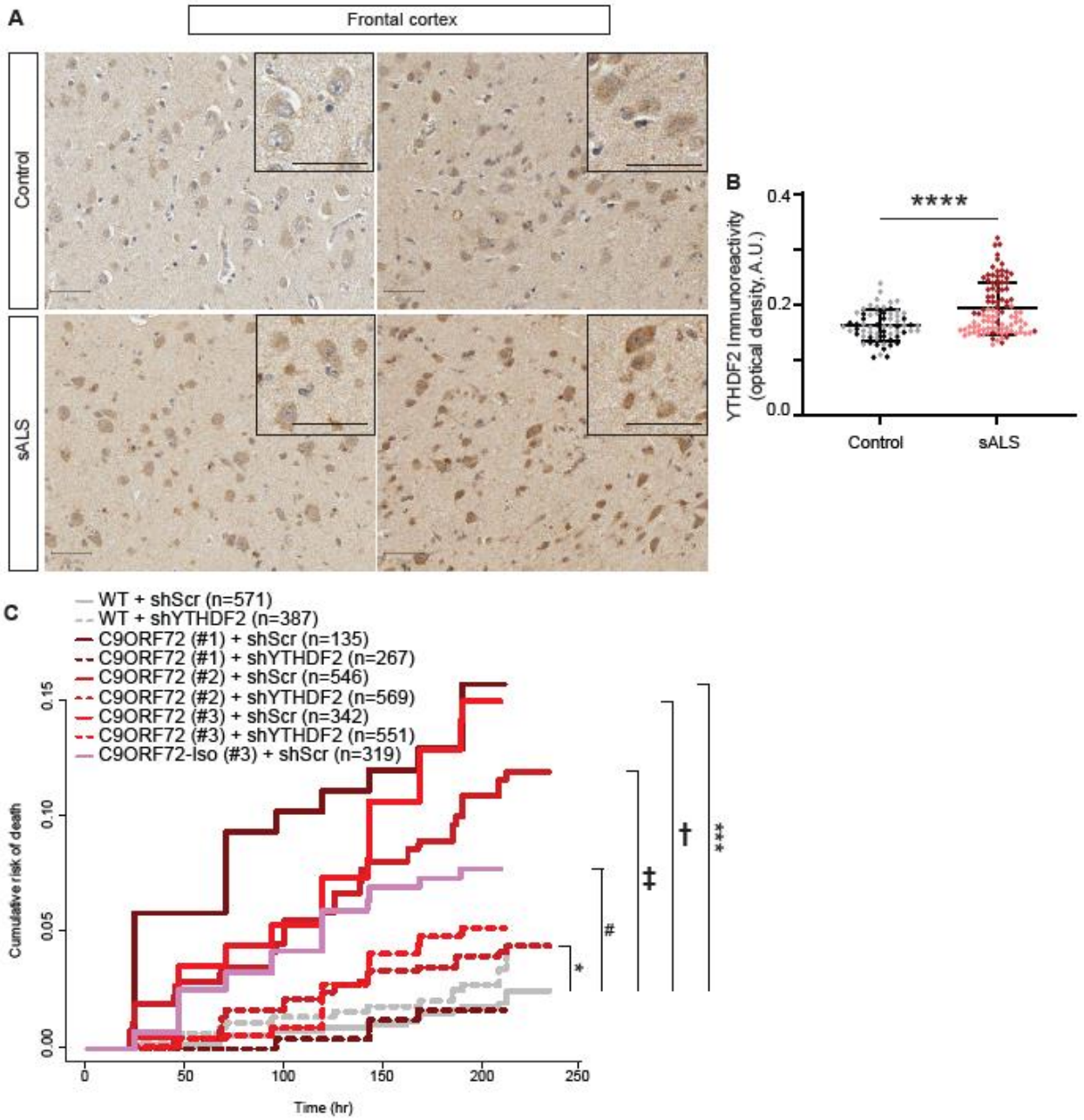
(**A**) Knockout of m6A writers Mettl3 (***p< 2.0 x$10^{-16}$, HR= 0=2.08; #p= 0.073, HR= 1.06; **p= 1.45x$10^{-3}$, HR= 1.12), Mettl14 (**B**, ***p< 2.0 x$10^{-16}$, HR= 2.27; *p= 0.025, HR= 1.08), or Wtap (**C**, ***p< 2.0 x$10^{-16}$, HR= 2.08, *p= 0.036, HR= 1.13) does not significantly

increase TDP43 dependent toxicity. Knockout of m6A eraser Fto (**D**, ***p<2.0 $x10^{-16}$, HR= 2.0; †p= $1.31x10^{-8}$, HR= 1.25) does not significantly increase toxicity, but knockout of Alkbh5 (**E,** ***p< 2.0 $x10^{-16}$, HR= 2.0; †p= $3.86x10^{-5}$, HR= 1.23; **p= $2.87x10^{-3}$, HR= 1.19) significantly increases toxicity. Knockout of m6A reader proteins Ythdf1(**F**, ***p<2.0 $x10^{-16}$, HR= 2.70; †p=1.02 $x10^{-4}$, HR= 1.15; ‡p= 3.49 $x10^{-6}$, HR= 1.21 ) and Ythdf3 (**H**, ***p<2.0 $x10^{-16}$, HR= 2.08; **p= $5.52x10^{-3}$, HR= 1.11; †p= 5.11 $x10^{-3}$, HR= 1.14) enhances toxicity, while Ythdf2 knockout (**G**, ***p<2.0 $x10^{-16}$, HR= 2.22; †p= 0.081, HR= 0.95; ‡p= $1.18x10^{-6}$, HR= 0.84) trends towards neuroprotection. ns= not significant. Statistical significance determined by Cox proportional hazards, with a minimum of 3 biological replicates per condition.




**Supplemental Figure 2.7: Knockout of additional m6A components at low levels of TDP43 expression is toxic.**

(**A**) Knockout of m6A writers Mettl3 (***p= 3.42 x$10^{-6}$, HR= 1.33; #p= 0.056, HR= 1.15); or Mettl14 (**B**, ***p< 2.00 x$10^{-16}$, HR= 1.67;) does not significantly increase toxicity, but Wtap knockout (**C**, ***p= 2.73 x$10^{-6}$, HR= 1.65; *p= 0.026, HR= 1.30) does increase TDP43-dependent toxicity. Knockout of m6A eraser Fto (**D,** ***p= 6.74 x$10^{-8}$, HR= 1.37) does not significantly increase toxicity despite Alkbh5 (**Fig. 5G**) significantly increasing TDP43-mediated toxicity. Knockout of m6A readers Ythdf1 (**E**, ***p<2.0 x$10^{-16}$, HR= 1.76; *p= 0.010, HR= 1.23) and Ythdf3 (**F**, ***p<7.18 x$10^{-8}$, HR= 1.47; *p= 0.026, HR= 1.22) enhance toxicity. ns= not significant. Statistical significance determined by Cox proportional hazards, with a minimum of 3 biological replicates per condition.




**Supplemental Figure 2.8: YTHDF2 immunoreactivity in frontal cortex.**
(**A**) Immunostaining of YTHDF2 in control and sALS patient frontal cortex samples. Scale bar= 50µm. (**B**) Quantification of YTHDF2 immunoreactivity in frontal cortex neurons from control (n= 65) and sALS (n= 99) samples. ****p<0.0001, Mann-Whitney test. (**C**) Knockout of YTHDF2 in *C9ORF72* iNeurons reduces toxicity (***p = $1.33 \times 10^{-6}$, HR= 7.36; †p= $6.74 \times 10^{-9}$, HR= 5.21; ‡p= $1.16 \times 10^{-3}$, HR= 6.90; #p= 0.045, HR= 1.9. *p= 0.031, HR= 3.66). Statistical significance determined by Cox proportional hazards.

**Table 2.1: Post-mortem samples**

| Diagnosis | Age | Sex | Experiment |
|---|---|---|---|
| Control 1 | 66 | M | m6A array |
| Control 2 | 85 | M | m6A array |
| Control 3 | 68 | M | m6A array |
| sALS 1 | 65 | M | m6A array |
| sALS 2 | 64 | M | m6A array |
| sALS 3 | 65 | M | m6A array |
| sALS 4 | 81 | M | m6A array |
| Control 4 | 76 | F | m6A IHC |
| Control 5 | 88 | M | m6A IHC, YTHDF2 IHC |
| Control 6 | 56 | M | m6A IHC, YTHDF2 IHC |
| ALS 1 | 66 | F | m6A IHC |
| ALS 2 | 64 | M | m6A IHC |
| ALS 3 | 68 | M | m6A IHC |
| Control 7 | 48 | M | YTHDF2 IHC |
| ALS 4 | 76 | M | YTHDF2 IHC |
| ALS 4 | 60 | F | YTHDF2 IHC |
| ALS 6 | 99 | M | YTHDF2 IHC |

**Table 2.2: Plasmids and primers used to create plasmids**

| Construct | Source | Amplicon/ Insert/Target | Sequence (5' to 3') |
|---|---|---|---|
| pGW1-EGFP[32,222] | | | |
| pGW1-mApple[32,222] | | | |
| pGW1-TDP43(WT)-EGFP[32,222] | | | |
| pGW1-Halo[217] | | | |
| pGW1-TDP43-TEV-Halo[21] | | | |
| pCMV-APOBEC1-YTH[104] | | | |
| pCMV-APOBEC1-YTH[104] | | | |
| pCaggs-mCherry-TBR[21,216,217] | | | |
| pE-6xHis-SUMO-TDP43(WT) | Dr. James Shorter | | |
| *CFTR* minigene[256] | | | |
| pSpCas9(BB)-2A-GFP | Addgene #48138 | | |
| pcDNA-flag-YTHDF2 | Addgene #52300 | | |
| pcDNA-flag-METTL3 | Addgene #53739 | | |
| pcDNA-flag-METTL14 | Addgene #53740 | | |
| GIPZ Non-silencing Lentiviral shRNA Control | Horizon Discovery #RHS4346 | Non-targeting | |
| GIPZ Lentiviral Human YTHDF2 shRNA | Horizon Discovery #RHS4430-200182983 | YTHDF2 | |

| | | | |
|---|---|---|---|
| pCaggs-mCherry-mTBR | This paper | TARDBP 3' UTR, SDM A-G | F: CATTATGCACCACCAAGCCTCTG CAC GCGCTCTC |
| | | | R: GCTTTGCAGGAGGGCTTGAAGCA GAG |
| pGW1-YTHDF2-Halo | This paper | YTHDF2 | F: gcgACCGGTATGGATTACAAGGAC G |
| | | | R: cgcTCTAGATTTCCCACGACC |
| pGW1-YTHDF2-2A-GFP | This paper | YTHDF2 | F: ACGCGGTACCCCATGTCGGCCA GCAG CCTC |
| | | | R: GCACGTCGACTTATTTCCCACGA CCTTG |
| pSpCas9(BB)-2A-GFP + NeuN sgRNA | This paper | NeuN | F: CACCGACCGTCTGGGTCCCAGC GAT |
| | | | R: AAACATCGCTGGGACCCAGACG GTC |
| pSpCas9(BB)-2A-GFP + Mettl3 sgRNA | This paper | Mettl3 | F: CACCGGCTGGGCTTAGGGCCAC TAG |
| | | | R: AAACCTAGTGGCCCTAAGCCCAG CC |
| pSpCas9(BB)-2A-GFP + Mettl14 sgRNA | This paper | Mettl14 | F: CACCGGATTCTTCTGGAGCCTCC TC |
| | | | R: AAACGAGGAGGCTCCAGAAGAAT CC |
| pSpCas9(BB)-2A-GFP + Wtap sgRNA | This paper | Wtap | F: CACCGGCCGCCAGTCACACAGG CCG |
| | | | R: AAACCGGCCTGTGTGACTGGCG GCC |
| pSpCas9(BB)-2A-GFP + Fto sgRNA | This paper | Fto | F: CACCGGCTGCACAAAGAGGTCC CCG |
| | | | R: AAACCGGGGACCTCTTTGTGCAG CG |

| | | | |
|---|---|---|---|
| pSpCas9(BB)-2A-GFP + Alkbh5 sgRNA | This paper | Alkbh5 | F: CACCGGCCTGCCTTGTAGTTGTCCC |
| | | | R: AAACGGGACAACTACAAGGCAGGCC |
| pSpCas9(BB)-2A-GFP + Ythdf1 sgRNA | This paper | Ythdf1 | F: CACCGGCTGTTTTTGGGCAACCTGG |
| | | | R: AAACCCAGGTTGCCCAAAAACAGCC |
| pSpCas9(BB)-2A-GFP + Ythdf2 sgRNA | This paper | Ythdf2 | F: CACCGGCTGTAGTAACTGGGTAAGT |
| | | | R: AAACACTTACCCAGTTACTACAGCC |
| pSpCas9(BB)-2A-GFP + Ythdf3 sgRNA | This paper | Ythdf3 | F: CACCGGCTCTCCCAAGAGAACTAGG |
| | | | R: AAACCCTAGTTCTCTTGGGAGAGCC |
| pSpCas9(BB)-2A-GFP + Atxn2 sgRNA | | | F: CACCGCAGCAGTTCTCGAGGAGGG |
| | | | R: AAACCCCTCCTCGAGAACTGCTGC |
| | | | |

**Table 2.3: Antibodies**

| Target | Source | Catalog number | Antibody registry number | Species | Dilution |
|---|---|---|---|---|---|
| TDP43 | Proteintech | 10782-2-AP | AB_2892214 | Rabbit | 1:100 |
| NeuN | Cell Signaling Technologies | 24307T | AB_2651140 | Rabbit | 1:500 |
| YTHDF2 | Proteintech | 24744-1-AP | AB_2687435 | Rabbit | 1:100 |
| m6A | Millipore Sigma | ABE572 | AB_2892213 | Rabbit | 1:500 |
| m6A | Cell Signaling Technologies | 56593S | AB_2799515 | Rabbit | 1:500 |
| m6A | Millipore Sigma | ABE572-I-100UG | AB_2892214 | Rabbit | 1:500 |
| HaloTag ligand (JF646 dye) | Promega | GA1120 | N/A | N/A | 1:20,000 |
| Anti-rabbit Alexa Fluor 488 | ThermoFisher | A-11034 | N/A | Goat | 1:250 |

| | | | | | |
|---|---|---|---|---|---|
| (secondary antibody) | | | | | |
| Anti-rabbit HRP (secondary antibody) | Jackson ImmunoResearch Labs Inc. | 111-035-003 | N/A | Goat | 1:10,000 |

**Table 2.4: Oligonucleotides**

| Target | Source | Sequence (5' to 3') |
|---|---|---|
| mCherry | This paper | F: ATGGTGAGCAAGGGCGAGGA |
| | | R: GATCTCGAACTCGTGGCC |
| *CFTR* | Ayala *et al.*, 2006[256] | F: caacttcaagctcctaagccactgc |
| | | R: taggatccggtcaccaggaagttggttaaatca |
| Clip34nt_unmodified | IDT | 5IR800CWN-AGAGACUUGGUGGU |
| Clip34nt_m6A | IDT | 5IR800CWN-AGAG/iN6Me-A/CUUGGUGGU |
| N_TARDBP_F | IDT | CACCGGAAATACCATCGGAAGACGA |
| N_TARDBP_R | IDT | CACCGGGGCTCATCGTTCTCATCTT |

**Table 2.5: Cell lines**

| Name | Source | Identifier |
|---|---|---|
| HEK293T | ATCC | CRL-3216 |
| Primary rat neurons | University of Michigan Unit for Laboratory Animal Medicine | |
| Human iPSCs | University of Michigan ALS Repository | 1021, 0883, 0312 |
| | Cedars Sinai iPSC Repository | Cs29i, Cs29 |

**Table 2.6: Human iPSC lines**

| Line | Identifier | Gene | Mutation | Sex | Age at biopsy | Age at onset | Tag | Integrated cassette |
|---|---|---|---|---|---|---|---|---|
| 1021 | WT | *TARDBP* | - | F | 54 | - | TDP43-Dendra 2 (C) | CLYBL:TO-hNgn1/2[217] |
| 1021 | M337V | *TARDBP* | Isogenic M337V[229] | F | 54 | - | TDP43-Dendra 2 (C) | CLYBL:TO-hNgn1/2[217] |
| 312 | C9ORF72 (#1) | *C9ORF72* | HRE | M | 52 | 54 | - | CLYBL:TO-hNgn1/2[217] |
| 883 | C9ORF72 (#2) | *C9ORF72* | HRE | M | 49 | 51 | - | CLYBL:TO-hNgn1/2[217] |
| Cs29i | C9ORF72 (#3) | *C9ORF72* | HRE | M | 47 | NA | - | CLYBL:TO-hNgn1/2[217] |
| Cs29i | C9ORF72 ISO (#3) | *C9ORF72* | Isogenic corrected[260] | M | 47 | NA | - | CLYBL:TO-hNgn1/2[217] |

NA: not assessed; HRE, hexanucleotide repeat expansion mutation; (C), carboxyl terminus

# Chapter 3: Discussion and Future Directions

## 3.1 Overview

In Chapter 2 of this dissertation, we demonstrate that TDP43 binds to methylated RNAs, and the large majority of transcripts bound by TDP43 contain m6A. In identifying the prevalence of m6A in TDP43 substrates, we found a spatial relationship between m6A sites in the 3' UTR and UG rich regions that TDP43 binds to. We detected an m6A site in the 3' UTR of the *TARDBP* transcript itself, specifically within a region that TDP43 was shown to bind. Removal of this m6A site diminished TDP43 binding, suggesting that m6A is crucial for TDP43 RNA substrate recognition and autoregulation. A hallmark of ALS is TDP43 mislocalization and cytoplasmic aggregation and given our observations of TDP43 recognizing methylated RNA and m6A impacting autoregulation, we measured m6A changes in sALS and control spinal cord. We observed significant hypermethylation of both mRNA and lncRNA from disease samples compared to controls, indicating that that RNA homeostasis may be disrupted in ALS because of the abundance of RNA modifications. This observation prompted investigation into the m6A pathway, utilizing CRISPR to knockout m6A readers, writers, and erasers in an attempt to modulate TDP43 toxicity. While knocking out most of the genes led to increased toxicity, only the m6A reader YTHDF2 reduced TDP43-mediated toxicity in primary rodent and human iNeuron models. Together with hypermethylation in ALS spinal cord, we also observed increased levels of YTHDF2. These findings allude to the importance of m6A for TDP43 binding and disease pathogenesis. However, these findings evoke

further questions regarding the contribution of m6A to TDP43 function, and if m6A is responsible or concurrent with TDP43 pathology in ALS. This chapter will discuss further questions and potential next steps for interrogating the role of m6A in TDP43 pathology and disease.

### 3.2 RNA stability

As previously described, when investigating stability in ALS models and with TDP43 overexpression, we observed global RNA instability specifically within transcripts pertaining to oxidative phosphorylation enzymes and ribosomal proteins[1]. However, in Mettl14 conditional knock out mice, these groups of transcripts are most upregulated[2]. These findings suggest that m6A may influence the stability of transcripts that are destabilized in ALS. The predominant effect of adding m6A to an RNA substrate is to destabilize it, where methylation decreases half-life of that RNA and increasing the amount of m6A per transcript further destabilizes it[3–5]. This is paired with exon and UTR length, as transcripts with shorter exons or UTRs tend to have less methylation than longer RNAs[6,7]. m6A levels are variable across different tissue and cell types, and cells within the nervous system contain the most methylation, perhaps as a mechanism of dynamic regulation[8–11]. Furthermore, m6A increases with age and altered m6A levels have contributed to the aging process and age-related diseases[12–14].

The data presented in this thesis argue that m6A could be influencing the stability of TDP43 RNA substrates. Our DART-seq data indicate that the majority (<90%) of transcripts bound by TDP43 contained m6A marks compared to the >10% that did not. The methylation profile differed from other published data sets however, where we saw high levels of methylation within the coding sequence and slightly higher than in the 3’

UTR. Typically, m6A is enriched at the end of the coding sequence and beginning of 3' UTR before it decreases moving down the UTR. We attribute the differences to using TDP43 to isolate RNA targets, instead of the global transcriptome that was measured in other data sets. Slightly complicating this finding is the relative absence of m6A from intronic regions[15–17]. We similarly observed that intronic regions lacked m6A, which suggests that m6A may not be critical for TDP43 to bind within introns or that intronic binding of TDP43 may not influence the stability in all transcripts. Instead, TDP43 binding in intronic regions could facilitate additional protein interactions that may influence the RNAs fate. However, we identified a relationship where within 20 nt of an m6A site there was a TDP43 binding motif in the 3' UTR of targets regulated by TDP43. These results suggest that m6A sites in the 3' UTR of TDP43 regulated genes may be functionally relevant.

To investigate the potential relevance of these m6A sites, I propose experiments that modulate m6A levels in cells to see if there are changes in the abundance or stability of target RNAs. These experiments may be performed in HEK293 cells for initial proof-of-principle, but differences in m6A levels and location depending on the cell type experiments should be performed in the closest to human systems available, like patient derived induced pluripotent stem cells (iPSCs). A way to measure RNA stability for this approach is with Roadblock qPCR[18]. Here, cells are metabolically labeled with a nucleoside analog 4-thiouridine (4sU), then the RNA is extracted and treated with a thiol reactive compound N-ethylmaleimide (NEM) that modifies 4sU and sterically interferes with reverse transcription of 4sU-containing transcripts during cDNA production before analyzing with qPCR. This approach allows for determination of transcript half-life

without stopping transcription, getting a more accurate measurement of RNA half-life. In these experiments, I would hypothesize that RNA stability is decreased with increased methylation and therefore target RNAs have a decreased half-life.

Alternatively, there are other approaches for replicating hypermethylation to study how methylation influences RNA stability. Expressing METTL3/14 is sufficient to increase methylation *in vitro*[19]. In an attempt to drive hypermethylation, lentiviral expression of METTL3/14 in iNeurons could potentially raise the levels of methylation to ask how increased m6A changes RNA stability. Given that overexpression may not improve methylation levels to the degree needed to see a change in RNA stability, the opposite approach could be taken to knockdown or knockout METTL3/14. METTL3/14 are the main writers of methylation and knockout of these two enzymes is shown to reduce m6A levels[20]. Under environments of overexpression or knockout of METTL3/14, qPCR can be performed to measure the abundance of TDP43 substrates to see if their levels are regulated by methylation, or the stability can be measured by Roadblock qPCR. Importantly, validation of METTL3/14 knockout determined by qPCR or Western blot would confirm knockout, and measuring m6A abundance by dot blot provides a simple way to measure the expected m6A reduction.

Another way that m6A can influence RNA stability is during transcription, and there is sufficient evidence that supports m6A deposition occurring co-transcriptionally. The enhancer protein CEBPZ recruits METTL3 to promoter regions to induce methylation within protein-coding regions of transcripts. METTL3 also is recruited to chromatin in a transcription-dependent manner to methylate nascent chromatin associated RNA (chrRNA) transcripts[7,21,22]. METTL14 can recognize H3K36me3

marking active genes, indicating that methylation components are recognizing actively transcribed genes to methylate the produced RNA[23]. Converting A into m6A co-transcriptionally is dependent on the activity of RNA polymerase II (RNA Pol II)[24]. Slowing down, but not stopping RNA Pol II elongation increases methylation levels, since the methylation complex has more time to act on the RNA being transcribed[24]. Transcripts that accrued more methylation because of slowed RNA Pol II also had a lower translation efficiency, perhaps due to destabilized RNA. This study treated HeLa cells with camptothecin at low doses to restrict RNA Pol II speed and artificially increase methylation levels. This approach serves as a viable alternative to overexpression systems and can be applicable to various cell types, although somewhat artificial because global interference with translation could obscure m6A specific phenotypes.

Additionally, we found hypermethylation of transcripts compared to controls in sALS spinal cord. Comparing methylated transcripts to the DART-seq data we found an overlap of 90%, suggesting that both approaches yielded similar results. Among the hypermethylated transcripts, we identified several that were related to ALS and other neurodegenerative diseases (*CSNK1E, TIA1, hnRNPAB, ANXA11, PARK7, WDR45*) and found over three hundred transcripts that appear to be regulated by TDP43. If the main consequence of m6A is to destabilize RNA, then we can infer that there is widespread destabilization of RNA which would match previous findings from our lab[1]. However, we do not know if there is a relationship between TDP43 mislocalization and hypermethylation in specific cells, or if one precedes the other.

To probe for this answer, we can perform immunohistochemistry (IHC) on spinal cord sections using antibodies against TDP43 and m6A. This dual-labeling approach

will allow us to visualize TDP43 mislocalization in spinal motor neurons and determine if m6A is upregulated or downregulated in those cells. If successful, this method should also be applied to familial forms of ALS to see if hypermethylation is a conserved feature of ALS, which could offer a new shared pathway to interrogate for therapeutic development. Additional future experiments will be needed to assess how changes in methylation in spinal cord impact RNA stability but could provide a novel approach to treating ALS patients.

While the m6A array did not discover any novel m6A sites, it informed of changes related to the overall amount of m6A in RNAs produces by a gene. Many of the m6A RNA-seq experiments are only capable of detecting m6A sites within transcripts, not measuring changes in methylation across transcripts and isoforms. The absolute fraction of RNAs in which m6A sites are methylated is referred to as m6A stoichiometry, and this percentage greatly influences the fate of RNAs[25,26]. For example, if a methylated RNA causes destabilization, 1% of transcripts being methylated will likely not cause a drastic effect. However, if that number jumps to 30%, then a much larger proportion of transcripts will be destabilized, potentially leading to negative downstream consequences[25]. Since m6A alters secondary structure, methylation at one site may have a completely different structure than that of a similar isoform without m6A. These differences are sufficient to recruit different RNA binding proteins and guide the fate of the two isoforms in different directions[26]. Despite recent advances in methods for detecting and measuring m6A, quantifying m6A stoichiometry remains a difficult but much needed future ability.

### 3.3 TDP43 autoregulation

TDP43 regulates its protein levels by binding to the TDP43 binding region (TBR) located in the 3' UTR of its own transcript[27–30]. Binding to this region triggers alternative splicing resulting in an unstable transcript and resulting decreased protein expression[27,28,31]. TDP43 is an essential protein, and regulation of its protein levels are critical to stay in homeostasis and prevent cell death due to TDP43-mediated toxicity. In ALS, there are various forms of TDP43 pathology that indicate dysregulation such as cytoplasmic mislocalization, nuclear clearance, and aggregation. Some ALS cases present with *TARDBP* mutations in the 3' UTR, potentially because the mutations might interfere with autoregulation given the importance of the TBR to TDP43 autoregulation[32]. As such, it is vital to understand the mechanisms of TDP43 autoregulation to prevent future dysregulation leading to disease.

In Chapter 2 we used DART-seq to find m6A sites throughout the transcriptome in an antibody-independent manner. We located an m6A site within the TBR, specifically in the sequence determined to be sufficient for TDP43 binding[30]. Methylation at this site appears to be critical because when the DRACH motif was altered in a reporter of the *TARDBP* 3' UTR, TDP43 exhibited reduced binding to the mutant reporter compared to the wild-type. This reporter contains an mCherry tag upstream of the TBR, and when TDP43 binds to the reporter there is a loss of fluorescence, likely due to a combination of nuclear retention of the spliced transcript and nonsense mediated mRNA decay[33]. In rat primary neurons, we noticed a significant change in fluorescence with TDP43 binding to the wild-type reporter that was not seen in the mutant. Moreover, the total fluorescence was much lower for the mutant reporter

compared to the wild type. These findings suggest that TDP43 requires m6A to recognize its own transcript, and for efficient regulation of its transcript levels. Despite the results from these studies, EMSAs using a small probe of the TBR produced a conflicting result in that the presence of m6A slightly deterred binding. Differences between both results may signify that there are other components involved or structural aspects with the reporter assay in cells that are not present in the *in vitro* EMSAs.

Identifying the m6A site within the TDP43 binding region of the *TARDBP* transcript could be an important step to understanding the molecular mechanisms that govern autoregulation. However, further investigation is needed to determine if this site is conserved across cell types, specifically within neurons and in ALS. DART-seq can be performed *in vitro* after purification of the APOBEC1-YTH fusion protein to identify m6A sites in RNA from iNeurons or patient RNA. Using this method, m6A sites in any sample of RNA can be identified across the transcriptome, or specific targets can be isolated via Sanger sequencing.

Previously, we only measured how loss of an m6A site would influence TDP43 binding. We hypothesized that TDP43 binds methylated RNAs, but did not look at how adding m6A sites would change autoregulation. Conversely, we can attempt to measure how introducing methylation sites would affect TDP43 binding by adding DRACH sites to the reporter construct. Unfortunately, this would add more DRACH sites to the reporter, but these sites would not necessarily be methylated. If the sites are methylated, then we predict that TDP43 would bind to the reporter more readily and would expect higher levels pulled down by TDP43. Increased TDP43 binding due to more m6A on the reporter would also decrease fluorescence levels, and should

destabilize the reporter more compared to controls. To confirm the presence of m6A on the newly inserted DRACH motifs, we could employ DART-seq APOBEC1-YTH to *in vitro* deaminate the obligate C following m6A sites in RNA derived from overexpression of the mutant reporter. Alternatively, use of bacterial enzyme MazF could determine presence of m6A[34]. MazF cleaves ACA motifs, but if m6A is present, the methylation prevents cleavage. This method offers another relatively simple antibody-independent method to determine if m6A is present at specific sites. However, this only works if the m6A site has an ACA motif so design of the inserted DRACH is important for the success of this experiment.

Prior work has demonstrated that m6A marks are enriched in the 3' UTR and near the stop codon, suggesting that m6A may be affecting alternative polyadenylation[8,35]. Alternative polyadenylation is a mechanism of gene regulation in which different 3' ends of a transcript are produced by RNA Pol II[36]. The *TARDBP* transcript produces many isoforms as a result of alternative polyadenylation, with most differences located in the exon 6 and the 3' UTR[33,37]. Previous work from our lab determined that an alternatively spliced *TARDBP* transcript produces a shortened TDP43 isoform (sTDP43) with a novel 18 amino acid tail[38]. Expression of sTDP43 was tied to neuronal activity, capable of mislocalizing TDP43 to the cytoplasm, and was elevated in ALS lumbar spinal cord[38]. The m6A writer VIRMA binds polyadenylation factors in an RNA dependent manner, and depletion of VIRMA led to increased 3' UTR length likely due to impaired polyadenylation resulting from less methylation[39]. With evidence of hypermethylation in ALS spinal cord samples, it is conceivable that more methylation led to alternative polyadenylation creating new shortened isoforms of

TDP43. To confirm this theory, future experiments should seek to understand how specific deposition of m6A within the *TARDBP* transcript influence isoform production.

### 3.4 Modulating TDP43 toxicity and ALS pathology

ALS is characterized by the loss of motor neurons, affecting the brain and spinal cord[40]. Over time, loss of functional neurons results in weakness in limbs, muscle atrophy, and paralysis with cause of death most often being respiratory failure within 3-5 years after diagnosis. ALS is phenotypically variable, and mutations in several genes have been shown to cause disease[41]. However, inherited mutations only account for around 10% of ALS cases, whereas most are sporadic. Although, a common unifying factor across almost all ALS patients is evidence of TDP43 pathology[42].

In an effort to understand if modulation of the m6A pathway could rescue TDP43-mediated toxicity, we performed a candidate CRISPR based screen knocking out m6A writers, erasers, and readers in primary rat neurons. In doing so, most knockouts increased TDP43 toxicity but knockdown of m6A reader YTHDF2 reduced TDP43-mediated toxicity. Encouraged by this result, we applied lentiviral delivered shRNAs targeting YTHDF2 in iNeuron models of ALS and similarly saw that YTHDF2 knockdown decreased TDP43 toxicity. YTHDF2 overexpression itself is toxic and in ALS patient spinal cord we observed elevated levels of YTHDF2. These data suggest that YTHDF2 may be a potential therapeutic target for treating ALS.

However, why YTHDF2 knockout rescues TDP43 toxicity and its what causes YTHDF2 upregulation is not known. In a large scale mass spectrometry based pulldown experiment to identify TDP43 targets, YTHDF2 weakly interacts with TDP43[43]. Despite this early evidence of an interaction, in available RNA-seq and proteomics data sets

from ALS tissue and models, there is no further evidence that TDP43 and YTHDF2 interact. Additionally, little is known about potential single nucleotide polymorphisms (SNPs) in YTHDF2 or any changes in allelic frequency in normal or ALS conditions. One study identified a polymorphism in the fourth intron of YTHDF2, located in a (TG)n microsatellite repeat[44]. In this study, the authors were seeking to identify polymorphisms associated with longevity and found that in 137 individuals over 100 years old 30% exhibited 15 repeats of TG ($(TG)_{15}$), where only 27% of 275 young individuals had $(TG)_{15}$[44]. Interestingly, the older individuals were nearly twice as likely (15.3% vs 8.0%) to be homozygous at this locus for the repea[44]t. The authors concluded that perhaps the 15 repeat length, and not the other expansions ranging from 12 to 27, provided some association to longevity[44]. However, this study was only performed within the Italian population, so there is no data on how common this allele is in the general population. This does provide an area to investigate in ALS to determine if there are any changes to the $(TG)_n$ repeat length that may be associated with disease or disease severity.

As mentioned in Chapter 1, YTHDF2 helps regulate RNA stability by recognizing methylated RNAs and recruiting the CCR4-NOT deadenylation complex to degrade them[45]. Given that we observed hypermethylation, its possible that YTHDF2 expression is upregulated in response to the increased amount of methylated RNAs. Further testing is needed to understand the effect that YTHDF2 knockout/knockdown has on RNA stability, specifically of TDP43 targets. Roadblock qPCR can be performed in ALS iNeurons at baseline and with YTHDF2 depletion to measure any changes in transcript stability. Additionally, a next step of attempting YTHDF2 reduction in animal models of ALS would be important for validating this phenotype. Our data show that YTHDF2

knockdown is capable of reducing toxicity in C9ORF72 and TARDBP mutant iPSC lines. Testing YTHDF2 knockdown in animal models that have similar mutations would be an appropriate place to start, to see if the phenotype carries through. Successful rescue of disease progression by reduced neuronal death or improved motor functions would be an important step towards a legitimate therapeutic target. However, there is a possibility that YTHDF2 knockdown does not rescue toxicity in these models because methylation differences across species may change which RNAs are methylated[46,47]. In addition, ALS mouse models are not fully representative of disease due to difficulties in displaying ALS phenotypes and pathology, complicating outcomes of experimental studies and translation into human patients[48].

### 3.5 Conclusion

ALS is a devastating disease, with no cure and very limited therapeutic options. There have been many recent advances in understanding the genetics and some mechanisms of disease, yet TDP43 functions and what causes its mislocalization in disease is still being understood. This thesis describes a previously unknown capability of TDP43 in terms of its RNA binding preferences and uncovered a new pathway to interrogate as a way to control TDP43 pathology. Even though m6A dysregulation has been tied to other diseases, its consequences are still being realized and focus should be put on how m6A contributes to pathogenesis of ALS and other neurodegenerative diseases. Hopefully, with continued exploration of this pathway we can uncover a new therapeutic option to treat ALS patients who desperately need it.

**References**


1. Tank EM, Figueroa-Romero C, Hinder LM, et al. Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat Commun*. 2018;9(1):2845. doi:10.1038/s41467-018-05049-z
2. Koranda JL, Dore L, Shi H, et al. Mettl14 Is Essential for Epitranscriptomic Regulation of Striatal Function and Learning. *Neuron*. 2018;99(2):283-292.e5. doi:10.1016/J.NEURON.2018.06.007
3. Sommer S, Lavi U, Darnell JE. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *J Mol Biol*. 1978;124(3):487-499. doi:10.1016/0022-2836(78)90183-3
4. Wang X, Lu Z, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. 2014;505(7481):117-120. doi:10.1038/nature12730
5. Schwartz S, Mumbach MR, Jovanovic M, et al. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep*. 2014;8(1):284-296. doi:10.1016/j.celrep.2014.05.048
6. Ke S, Alemu EA, Mertens C, et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev*. 2015;29(19):2037-2053. doi:10.1101/gad.269415.115
7. Ke S, Pandya-Jones A, Saito Y, et al. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev*. 2017;31(10). doi:10.1101/gad.301036.117
8. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 2012;149(7):1635-1646. doi:10.1016/j.cell.2012.05.003
9. Cui Q, Shi H, Ye P, et al. m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep*. 2017;18(11):2622. doi:10.1016/J.CELREP.2017.02.059
10. Li F, Yi Y, Miao Y, et al. N6-methyladenosine Modulates Nonsense-mediated mRNA Decay in Human Glioblastoma. *Cancer Res*. September 2019:canres.2868.2018. doi:10.1158/0008-5472.CAN-18-2868
11. Widagdo J, Zhao QY, Kempen MJ, et al. Experience-Dependent Accumulation of N6-Methyladenosine in the Prefrontal Cortex Is Associated with Memory Processes in Mice. *J Neurosci*. 2016;36(25):6771. doi:10.1523/JNEUROSCI.4053-15.2016
12. Shafik AM, Zhang F, Guo Z, et al. N6-methyladenosine dynamics in neurodevelopment and aging, and its potential role in Alzheimer's disease. *Genome Biol*. 2021;22(1):17. doi:10.1186/s13059-020-02249-z
13. Jiang L, Lin W, Zhang C, et al. Interaction of tau with HNRNPA2B1 and N 6-methyladenosine RNA mediates the progression of tauopathy. *Mol Cell*. 2021;81(20):4209-4227.e12. doi:10.1016/J.MOLCEL.2021.07.038
14. Sun J, Cheng B, Su Y, et al. The Potential Role of m6A RNA Methylation in the Aging Process and Aging-Associated Diseases. *Front Genet*. 2022;13:936. doi:10.3389/FGENE.2022.869950/BIBTEX
15. Ke S, Pandya-Jones A, Saito Y, et al. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic

turnover. *Genes Dev*. 2017;31(10):990-1006. doi:10.1101/GAD.301036.117/-/DC1
16. Louloupi A, Ntini E, Conrad T, Ørom UAV. Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell Rep*. 2018;23(12):3429-3437. doi:10.1016/J.CELREP.2018.05.077
17. Wei G, Almeida M, Pintacuda G, et al. Acute depletion of METTL3 implicates N6-methyladenosine in alternative intron/exon inclusion in the nascent transcriptome. *Genome Res*. 2021;31(8):1395-1408. doi:10.1101/GR.271635.120
18. Watson MJ, Park Y, Thoreen CC. Roadblock-qPCR: a simple and inexpensive strategy for targeted measurements of mRNA stability. *RNA*. 2021;27(3):335. doi:10.1261/RNA.076885.120
19. Liu J, Yue Y, Han D, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol*. 2014;10(2):93-95. doi:10.1038/nchembio.1432
20. Meyer KD, Jaffrey SR. Rethinking m6A Readers, Writers, and Erasers. *Annu Rev Cell Dev Biol*. 2017;33:319-342. doi:10.1146/annurev-cellbio-100616-060758
21. Barbieri I, Tzelepis K, Pandolfini L, et al. Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nat 2017 5527683*. 2017;552(7683):126-131. doi:10.1038/nature24678
22. Knuckles P, Carl SH, Musheev M, Niehrs C, Wenger A, Bühler M. RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. *Nat Struct Mol Biol*. 2017;24(7):561-569. doi:10.1038/NSMB.3419
23. Huang H, Weng H, Zhou K, et al. Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nat 2019 5677748*. 2019;567(7748):414-419. doi:10.1038/s41586-019-1016-7
24. Slobodin B, Han R, Calderone V, et al. Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell*. 2017;169(2):326. doi:10.1016/J.CELL.2017.03.031
25. Gilbert W V., Bell TA, Schaening C. Messenger RNA modifications – Form, distribution, and function. *Science*. 2016;352(6292):1408. doi:10.1126/SCIENCE.AAD8711
26. Lewis CJT, Pan T, Kalsotra A. RNA Modifications and Structures Cooperate to Guide RNA-protein Interactions. *Nat Rev Mol Cell Biol*. 2017;18(3):202. doi:10.1038/NRM.2016.163
27. Polymenidou M, Lagier-Tourenne C, Hutt KR, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14(4):459-468. doi:10.1038/nn.2779
28. Ayala YM, De Conti L, Avendaño-Vázquez SE, et al. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J*. 2011;30(2):277-288. doi:10.1038/emboj.2010.310
29. Tollervey JR, Curk T, Rogelj B, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci*. 2011;14(4):452-458. doi:10.1038/nn.2778
30. Bhardwaj A, Myers MP, Buratti E, Baralle FE. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res*. 2013;41(9):5062-5074. doi:10.1093/nar/gkt189

31. Xu Y-F, Gendron TF, Zhang Y-J, et al. Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J Neurosci*. 2010;30(32):10851-10859. doi:10.1523/JNEUROSCI.1630-10.2010
32. Pesiridis GS, Lee VM-Y, Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009;18(R2):R156-62. doi:10.1093/hmg/ddp303
33. Eréndira Avendaño-Vázquez S, Dhir A, Bembich S, Buratti E, Proudfoot N, Baralle FE. Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev*. 2012;26(15):1679-1684. doi:10.1101/GAD.194829.112
34. Zhang Z, Chen LQ, Zhao YL, et al. Single-base mapping of m6A by an antibody-independent method. *Sci Adv*. 2019;5(7):250-253. doi:10.1126/SCIADV.AAX0250
35. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. 2012;485(7397):201-206. doi:10.1038/nature11112
36. Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol 2016 181*. 2016;18(1):18-30. doi:10.1038/nrm.2016.116
37. D'Alton S, Altshuler M, Lewis J. Studies of alternative isoforms provide insight into TDP-43 autoregulation and pathogenesis. *RNA*. 2015;21(8):1419-1432. doi:10.1261/RNA.047647.114
38. Weskamp K, Tank EM, Miguez R, et al. Shortened TDP43 isoforms upregulated by neuronal hyperactivity drive TDP43 pathology in ALS. *J Clin Invest*. 2020;130(3):1139-1155. doi:10.1172/JCI130988
39. Yue Y, Liu J, Cui X, et al. VIRMA mediates preferential m6A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov*. 2018;4(1):10. doi:10.1038/s41421-018-0019-0
40. Charcot J., Joffory A. Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arcives Physiol Neurol Pathol*. 1869;2:744-754.
41. Chia R, Chiò A, Traynor BJ. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *Lancet Neurol*. 2018;17(1):94-102. doi:10.1016/S1474-4422(17)30401-5
42. Neumann M. Molecular neuropathology of TDP-43 proteinopathies. *Int J Mol Sci*. 2009;10(1):232-246. doi:10.3390/ijms10010232
43. Freibaum BD, Chitta RK, High AA, Taylor JP. Global Analysis of TDP-43 Interacting Proteins Reveals Strong Association with RNA Splicing and Translation Machinery. *J Proteome Res*. 2010;9(2):1104-1120. doi:10.1021/pr901076y
44. Cardelli M, Marchegiani F, Cavallone L, et al. A polymorphism of the YTHDF2 gene (1p35) located in an Alu-rich genomic domain is associated with human longevity. *J Gerontol A Biol Sci Med Sci*. 2006;61(6):547-556. doi:10.1093/GERONA/61.6.547
45. Du H, Zhao Y, He J, et al. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun*. 2016;7(1):12626. doi:10.1038/ncomms12626

46. Liu J, Li K, Cai J, et al. Landscape and Regulation of m6A and m6Am Methylome across Human and Mouse Tissues. *Mol Cell*. 2020;77(2):426-440.e6. doi:10.1016/J.MOLCEL.2019.09.032
47. Jiang X, Liu B, Nie Z, et al. The role of m6A modification in the biological functions and diseases. *Signal Transduct Target Ther 2021 61*. 2021;6(1):1-16. doi:10.1038/s41392-020-00450-x
48. De Giorgio F, Maduro C, Fisher EMC, Acevedo-Arozena A. Transgenic and physiological mouse models give insights into different aspects of amyotrophic lateral sclerosis. *DMM Dis Model Mech*. 2019;12(1). doi:10.1242/DMM.037424/3045

# Appendix A: Golden Gate CRISPR and Cell Line Creation

## A.1 Introduction

Genetic manipulation of cells to investigate specific functions of genes has been performed since the early days of molecular biology. More recently, the discovery and application of clustered regularly interspaced short palindromic repeats (CRISPR) has allowed for an efficient and largely specific way to knockout gene function[1,2]. CRISPR and CRISPR associated (Cas) genes play important roles in certain bacteria and archaea, helping them avoid and eliminate invading bacteriophages[2]. CRISPR sequences within the genome of prokaryotic organisms was integrated after previous infections, and they use those same sequences as a defense system to prevent additional infections from similar bacteriophages. The Cas9 enzyme recognizes CRISPR sequences within the DNA and cleaves sequences that are complimentary, thereby protecting the cell from invasion[2,3].

Specifically, this protection occurs through three mechanisms, but most studied is the type II mechanism. In this case, invading DNA from viruses or plasmids are cut up into small fragments and incorporated into a CRISRP locus surrounded by a series of short 20 nt repeats[4]. The locus is then transcribed producing small RNAs called CRISPR RNA (crRNA) which act as a guide for endonucleases that target invading DNA based on the sequence complementarity[4]. Cas9 is the sole enzyme required for gene silencing in the type II system and participates in processing the crRNAs and breaking down the target DNA[4,5]. For site specific recognition and cleavage, Cas9

needs to interact with both a crRNA and a separate trans-activating crRNA (tracrRNA) which is complimentary to the crRNA[4]. Cas9 also requires the presence of a short sequence (2-5nt) called the protospacer associated motif (PAM), immediately downstream of the crRNA complementary sequence[6]. If all these components are present, Cas9 will create a double stranded break (DSB) in the DNA, rendering the invading DNA fragmented and maintaining the cells survival.

In 2012, this biological process was identified as a unique and powerful molecular biology tool for gene editing[4]. By combining the function of tracrRNA and crRNA into one synthetic single guide RNA (sgRNA), and including Cas9 enzyme, this simplified system generated successful genetic alterations. In combination, the sgRNA targets the Cas9 enzyme to the specific gene target, where the Cas9 enzyme creates DSBs and activates the DNA damage repair pathway of non-homologous end joining (NHEJ)[7,8]. DNA repair using NHEJ results in insertions or deletions that disrupt target locus. Another Cas9 enzyme, Cas9D10A, induces single-stranded nicks in each DNA strand and avoids NHEJ when supplied with a donor template, proceeds through the high-fidelity homology directed repair pathway (HDR)[1,4,8,9]. In both cases, a piece of DNA with homology to both sides of the cut can be used to endogenously incorporate tags to create knock-in cell lines instead of knockout[10].

Applying these principals, Sakuma and colleagues set out to combine multiple CRISPR guides in one plasmid also encoding the Cas9 enzyme to create an all-in-one targeting vector[11]. Utilizing specific Type IIS enzymes, these plasmids allow for digestion and insertion of guides on one step, and incorporation of up to 7 guides in one vector, or 3 paired nickase sgRNAs, via Golden Gate cloning. Commonly with hard to

detect proteins, direct tagging with a fluorescent protein is very difficult because of the low expression. However, modification at one locus is highly correlated with modification at a separate locus, and we take advantage of this using Golden Gate cloning to insert a fluorescent protein in one locus to observe changes in another[1,12]. In Appendix A, I describe adaptations to the protocol for improved sgRNA cloning efficiency and illustrate the potential applications of using Golden Gate cloning to create knock-in stem cell lines.

## A.2 Results

### *A.2.1 Cloning of sgRNA Vectors*

To create CRISPR guides against specific genes, the now obsolete sgRNA design tool from the Feng Zhang lab was used, although other bioinformatic tools now exist that have the same capabilities[13–16]. Guides against the safe harbor locus *CLYBL* were initially designed and used in testing incorporation into a backbone vector containing Cas9 nickase, pX330A_D10A1x4, and up to four sgRNAs: one into the "backbone" vector (pX330A_D10A1x4), and three "insert" vectors (pX330S_2…4, Fig. A.2.1). The guides were designed to have matching overhangs after digestion of the backbone vector with BbsI enzyme, a Type IIS enzyme that cleaves in downstream of its recognition site meaning once the sgRNAs are ligated correctly in the vector, the enzyme sites are destroyed. This allows for simultaneous digestion and ligation to increase the speed and efficiency of cloning. Step 1 of this process was cloning each sgRNA into a separate vector, with one going into the backbone vector and the other three into the insert vectors. With each guide driven by a separate U6 promoter. After transforming, colonies were screened using colony PCR with the forward primer against

the U6 promoter each sgRNA is driven by and the reverse primer against the antisense of the sgRNA to confirm correct insertion producing a band at 100 bp. Once each sgRNA was cloned into its respective vector, then step 2 proceeded which was the digestion and ligation reaction to insert each sgRNA along with its U6 promoter from the insert vectors into the backbone vector using a different enzyme BsaI, which acts similarly as before. After transforming into bacteria again, colony PCR confirmed the presence of four sgRNAs in the backbone vector. An additional check of enzyme digestion that results in 500 bp increases per sgRNA validated the presence of four sgRNAs in the final vector.

### *A.2.2 Creating iPSC Knock-in cell lines*

When attempting to create a knockout stem cell line of lowly expressed proteins, it can be difficult to determine which cells were transfected if checking by fluorescence. Understanding that CRISPR modification at one locus is correlated with modification at only locus[17], we created Cas9D10A containing plasmids of two paired sgRNAs to target two separate gene loci: one locus being a safe harbor locus, while the other was the gene of interest. Using sgRNAs against the *CLYBL* safe harbor locus, we supplied a HDR containing a Ngn1/2 cassette, shown to be sufficient to differentiate iPSCs into iNeurons, tagged with iRFP[18]. The other sgRNAs targeted *LC3*, a central protein in autophagy, to knockout function[19]. After transfecting sgRNAs along with the HDR into iPSCs, fluorescence microscopy was used to screen colonies for transfected cells, removing non-transfected cells, enriching for fluorescent cells until a homogeneous population is reached.

## A.3 Discussion

This work was an adaptation of work performed by Sakuma and colleagues in which they created vectors capable of multiplexed CRISPR. After optimization, Step 1 cloning yielded a success rate of about 95%. The protocol for Step 1 cloning can be applied to other CRISPR sgRNA cloning vectors that rely upon Type IIS enzymes for high efficiency cloning. Step 2 cloning was somewhat more cumbersome, with a success rate of about 50%. The expected timeline to complete a multiplexed CRISPR vector is about 10 days, if every step is correct initially. This provides a fast and efficient way to target multiple genes for investigation of gene function after knockout. Further work may build upon this protocol to enhance the efficiency or speed in building a multiplex vector.

Transfection of complete multiplex plasmids into iPSCs with a supplied HDR was successful in producing fluorescent cells. Unfortunately, it did not enhance the fluorescence intensity for detecting LC3, but this could be a result of the low expression of LC3. However, through repeated selection enrichment was possible to create homogeneous stocks that were fluorescing well above background signal. Important next steps are to confirm whether or not knockout was successful. These vectors may still be hindered by the low expression levels of endogenous proteins at this stage. Differentiation into iNeurons, or treatment with rapamycin to stimulate autophagy or chloroquine to inhibit autophagy can be used to observe changes in autophagy in LC3 knockout cells. Failure for normal autophagic function would be indicative of disrupted LC3 function. Taken together, Golden Gate multiplex cloning is successful at creating

multiplexed vectors targeting different loci in iPSCs, and further work should be performed to validate knockout and any subsequent phenotypes.

## A.4 Methods

Golden Gate Cloning

Oligos complementary to gene targets were created using CRISPR sgRNA creator tool from the Feng Zhang lab. These oligos were phosphorylated and annealed together and inserted into backbone vector pX330A_D10A1x4 or insert vectors pX330S_2-4 (Addgene #1000000055) using BbsI (NEB #R3539S) digestion and simultaneously ligated with Quick Ligase (NEB #M2200S). Resulting vectors were pooled and digested with BsaI (NEB #R3535) while simultaneously ligated together using T4 Ligase (NEB #M0202S) and transformed into DH5α bacteria. Detailed protocol can be found below (A.6)

Colony PCR

Following transformation, colonies were screened for the presence of inserted oligo using a forward primer against the U6 promoter and the reverse primer being against the antisense of the most downstream sgRNA added using 2xGoTaq Green Mastermix according to manufactures protocol. Following PCR, samples were run on a 1% agarose gel and imaged using a UV imager.

Enzyme Test Digest

Plasmid DNA was prepared from liquid culture using the Qiagen Miniprep kit (Qiagen #27106). 1 µg of plasmid DNA was incubated for 3 hours with AflIII (NEB # R0541S) and KpnI (NEB # R3142S) at 37°C. Digested DNA was run on a 1% agarose gel and imaged using a UV imager.

iPSC Culture, Transfection, and Selection

In brief, iPSC lines were cultured in Essential 8 (E8) media (Gibco A1517001) on plates coated with vitronectin (Gibco A14700) diluted 1:100 in Mg2+/Ca2+ -free phosphate buffered saline (PBS, 178 Gibco 14190-144). Cells were passaged every 5d using 0.5 mM EDTA (Sigma E7889) followed by gentle trituration in E8 media with a P1000 pipette. All lines are verified mycoplasma-free on a monthly basis.

iPSCs were transfected with 2.5 µg of HDR DNA containing Ngn1/2 cassette with 400 bp flanking regions homologous to CLYBL(in pUC-minus, synthesized by Blue Heron) or 1.6 µg of Golden Gate using Lipofectamine Stem (Invitrogen #STEM00003). The following morning media was changed back to E8. Cells were screened for iRFP fluorescence and repeatedly selected until a homogenous population was reached.

## A.5 Acknowledgements

I would like to thank Dr. Elizabeth Tank for her technical support and advice for cloning and iPSC work. In addition, I would like to thank Drs. Tetsushi Sakuma and Takashi Yamamoto for their help providing technical support.

## A.6 Detailed Golden Gate Protocol

Step 1 of Golden Gate Assembly: Insertion of annealed oligonucleotides

1. Phosphorylate and anneal oligos by combining:
    i. 1 uL of 100 uM oligo 1 with 1 uL of 100 uM oligo 2
    ii. 1 uL of 10X T4 Ligation Buffer (NEB)
    iii. 6.5 uL dd$H_2O$
    iv. 0.5 uL T4 PNK (NEB)
    v. 10 uL total when combined
    b. Anneal in thermocycler:

i. 37°C- 30 min

ii. 95°C- 5 min

iii. decrease to 25°C at 5°C/min

iv. 4°C- ∞

2. Insert oligos into backbone vector and insert vectors

    a. Backbone vectors are pX330A-1x2 → pX330A1x7 and pX330A_D10A-1x2 →pX330A_D101x7. Insert vectors are pX330S-2…pX330S-7. pX330A-1x# coincides with the number of gRNAs that can be inserted into the backbone vector. Select one gRNA to be inserted into backbone vector. Select the appropriate number of pX330S-# vectors needed to add in remaining gRNA for final construct.

    b. Insert sgRNA into pX330A_D10-1x4 vector. Insert remaining sgRNAs into pX330S-2, pX330S-3, and pX330S-4, and so on, respectively. Prepare:

        i. 100 ng backbone vector or insert vector

        ii. 0.5 µL of 10uM annealed oligos (prepared at 100 µM)

        iii. 2 µL 10X T4 DNA ligase buffer

        iv. 1 µL BbsI-HF (aka BpiI) enzyme (NEB)

        v. 0.5 µL Quick ligase

        vi. H2O to volume

        vii. 20 µL total volume

    c. Combine master mix, vector, and oligos and put into thermocycler and run at:

        i. 37°C- 5 min, 16°C- 10 min.

        ii. Repeat 6x

        iii. 4°C- ∞

    d. Perform an additional digestion to confirm for correct insertion of sgRNA. If inserted correctly, there should be no more BbsI sites. If plasmid is cut it will be linearized and not be successfully transformed into bacteria:

        i. Create mix for each reaction of:

            1. 0.5 µL of CutSmart

            2. 0.25 µL BbsI

    e. Add to tubes and run thermocycler at:

        i. 37°C- 1 hr

ii. 80°C- 5 min
iii. 4°C- ∞

3. Transform product into competent *E. coli*:
    a. To transform the plasmid with the inserted oligonucleotides add
        i. 20 µL of product from previous step
        ii. 50 µL of DH5α
    b. Place on ice for 20 mins
    c. Put at 42°C for 30 sec
    d. Return to ice for 2 min
    e. Add 100 µL of LB to tubes and place at 37°C shaking for 1 hour
    f. Plate on 50 µg/mL Spec or Amp plates to grow colonies of transformed *E. coli*
    g. Check insertion by colony PCR, BbsI digest and sequencing
4. Colony PCR:
    a. Select colony from plate and place into reaction mix:
        i. 2x GoGreen Master Mix- 10 µL
        ii. U6_F 10 µM primer- 1 µL
        iii. AS 10 µM oligo- 1 µL
        iv. H2O- 8 µL
        v. Total volume-20 µL
    b. Place in thermocycler with following program:
        i. 95°C- 3 min
        ii. 95°C- 30 sec
        iii. 50°C- 30 sec
        iv. 72°C- 15 sec
        v. Repeat steps ii-iv 30x
        vi. 72°C- 10 min
        vii. 12°C- ∞
        viii. Run on 1% agarose gel for 30 mins at 120V
    c. Positive result should be a band at ~100 bp.

d. Inoculate 4 mL media with possible positive colony. Use 3 mL to miniprep to get DNA

5. BbsI digest to confirm elimination of BbsI site
    a. 2 µL DNA from transformation
    b. 2 µL Cutsmart buffer
    c. 1 µL BbsI
    d. 15 µL H2O
    e. Incubate in 37°C water bath for 30 mins.
    f. Run on 1% agarose gel. If DNA is cut, then oligo either isn't inserted or inserted incorrectly.
6. Sequencing
    a. Use U6_F primer for sequencing

Step 2 of Golden Gate Assembly: Combining Insert vectors with Backbone vector

1. To combine the two vectors, prepare the following mixture in PCR tubes after miniprepping from Step 1. Create 4 mL culture from step 1 colony and miniprep 3 mL to get DNA:
    a. 75 ng pX330A backbone vector(1x4)
    b. 150 ng each pX330S insert vectors (3 if using 1x4)
    c. 1 µL 10X T4 DNA ligase buffer
    d. 1 µL T4 ligase
    e. 1 µL BsaI-HF (NEB)
    f. H2O to volume
    g. 10 µL total volume
    h. Place in thermocyler and run at:
        i. 37°C- 10 min
        ii. 22°C- 10 min
            1. Repeat 25x
        iii. 4°C - ∞
    i. Perform additional digestion to confirm insertion and successful removal of BsaI sites:
        i. 2 µL CutSmart
        ii. 1 µL BsaI-HF

iii. Put in thermocycler for:

1. 37°C- 30 min
2. 80°C- 5 min
3. 4°C- ∞

j. Transform as before:

i. 10 µL product into 50 µL competent *E.coli*
ii. Place on ice for 20 mins
iii. Put at 42°C for 30 sec
iv. Return to ice for 2 mins
v. Add 100 µL of LB
vi. Place at 37°C for 1 hour
vii. Plate on LB + Amp plates
viii. Incubate plates overnight at 37°C
ix. Check alignment by colony PCR, BsaI digest, KpnI/AflIII digest, and sequencing

2. Colony PCR:

a. 2x GoGreen Master Mix- 10 µL
b. F-primer-1$^{st}$ position Sense 10 µM oligo - 1 µL
c. R-primer- 4$^{th}$ position Anti sense 10 uM oligo- 1 µL
d. H2O- 8 µL
e. Place in thermocycler and run at:

i. 95°C- 30 sec
ii. 95°C- 15 sec
iii. 62°C- 15 sec
iv. 72°C- 2 min
v. Repeat steps 2-4 27 times
vi. 72°C- 50 sec
vii. 4°C- ∞
viii. Run on 1% agarose gel for 30 min at 120V

ix. Positive result should be band at 2 kb. Each insert size is around 500 bp. Smear pattern in well where there is distinct band at 1650 bp can also indicate a possible successful combination

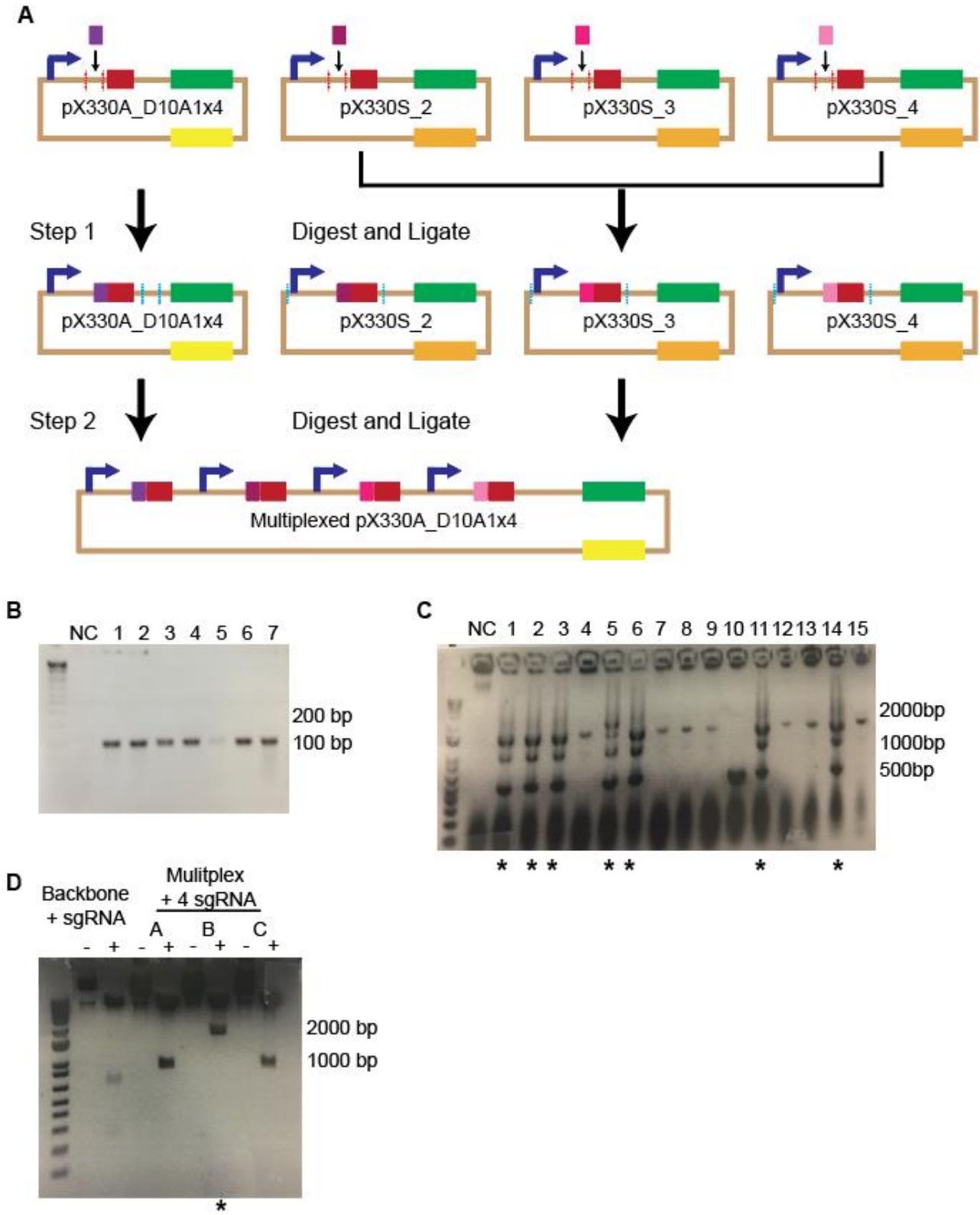
3. BsaI-HF digest:
    a. 2 µL DNA
    b. 2 µL Cutsmart
    c. 1 µL BsaI-HF
    d. 15 µL H2O
    e. Incubate 37°C for 30 mins. If four are aligned correctly, DNA should not be cut
4. KpnI/AflII Digest:
    a. 2 µL DNA
    b. 2 µL Cutsmart
    c. 1 µL KpnI-HF
    d. 1 µL AflIII
    e. 14 µL H2O
    f. Incubate in 37°C water bath for 2 hours. Cutsmart buffer is only 50% effective with both enzymes, needs to be run longer than other digests.
    g. Run on 1% agarose gel for 30 mins at 120V
    h. Positive results should be band at 2 kb. Enzymes cleave DNA right before first sgRNA and after last sgRNA
5. Sequencing:
    a. Use CRISPR-step2-F/R primers provided from Addgene kit described in table A.1

## References


1. Cong L, Ran FA, Cox D, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 2013;339(6121):819. doi:10.1126/SCIENCE.1231143
2. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science (80- )*. 2007;315(5819):1709-1712. doi:10.1126/SCIENCE.1138140/SUPPL_FILE/BARRANGOU.SOM.PDF
3. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science (80- )*. 2008;322(5909):1843-1845. doi:10.1126/SCIENCE.1165771/SUPPL_FILE/MARRAFFINI-SOM.PDF
4. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80- )*. 2012;337(6096):816-821. doi:10.1126/SCIENCE.1225829/SUPPL_FILE/JINEK.SM.PDF
5. Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nat 2011 4717340*. 2011;471(7340):602-607. doi:10.1038/nature09886
6. Swarts DC, Mosterd C, van Passel MWJ, Brouns SJJ. CRISPR Interference Directs Strand Specific Spacer Acquisition. *PLoS One*. 2012;7(4):e35888. doi:10.1371/JOURNAL.PONE.0035888
7. Overballe-Petersen S, Harms K, Orlando LAA, et al. Bacterial natural transformation by highly fragmented and damaged DNA. *Proc Natl Acad Sci U S A*. 2013;110(49):19860-19865. doi:10.1073/PNAS.1315278110/SUPPL_FILE/ST05.DOCX
8. Gong C, Bongiorno P, Martins A, et al. Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat Struct Mol Biol 2005 124*. 2005;12(4):304-312. doi:10.1038/nsmb915
9. Davis L, Maizels N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc Natl Acad Sci U S A*. 2014;111(10). doi:10.1073/PNAS.1400236111
10. Chen B, Gilbert LA, Cimini BA, et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*. 2013;155(7):1479. doi:10.1016/J.CELL.2013.12.001
11. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Reports 2014 41*. 2014;4(1):1-6. doi:10.1038/srep05400
12. Mali P, Yang L, Esvelt KM, et al. RNA-Guided Human Genome Engineering via Cas9. *Science*. 2013;339(6121):823. doi:10.1126/SCIENCE.1232033
13. Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods 2014 112*. 2014;11(2):122-123. doi:10.1038/nmeth.2812
14. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res*. 2014;42(W1):W401-W407. doi:10.1093/NAR/GKU410
15. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res*. 2016;44(W1):W272-W276. doi:10.1093/NAR/GKW398
16. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E.

CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*. 2019;47(W1):W171-W174. doi:10.1093/NAR/GKZ365
17. Yang H, Wu JJ, Tang T, Liu K De, Dai C. CRISPR/Cas9-mediated genome editing efficiently creates specific mutations at multiple loci using one sgRNA in Brassica napus. *Sci Rep*. 2017;7(1). doi:10.1038/S41598-017-07871-9
18. Cerbini T, Funahashi R, Luo Y, et al. Transcription Activator-Like Effector Nuclease (TALEN)-Mediated CLYBL Targeting Enables Enhanced Transgene Expression and One-Step Generation of Dual Reporter Human Induced Pluripotent Stem Cell (iPSC) and Neural Stem Cell (NSC) Lines. *PLoS One*. 2015;10(1). doi:10.1371/JOURNAL.PONE.0116032
19. Tanida I, Ueno T, Kominami E. LC3 and Autophagy. *Methods Mol Biol*. 2008;445:77-88. doi:10.1007/978-1-59745-157-4_4

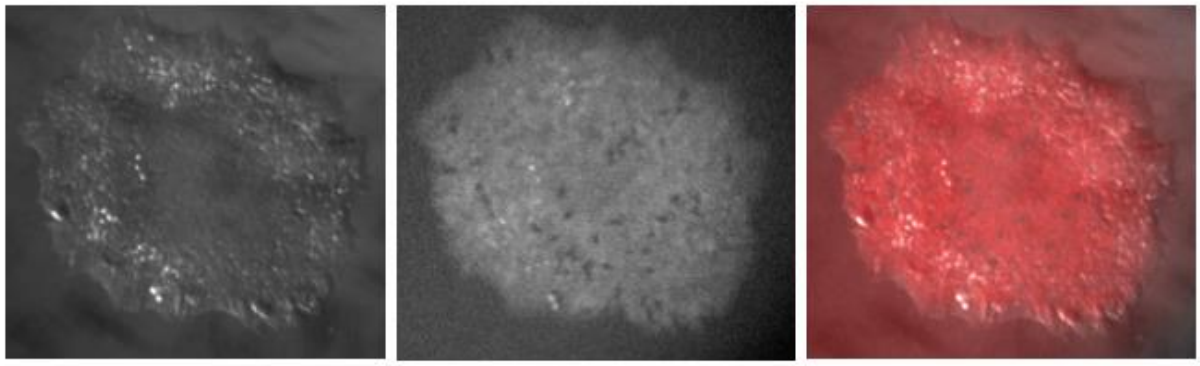
## Figures




**Appendix Figure A.1: Overview of Golden Gate CRISPR cloning.**

(A) Adapted from Sakuma *et al.*[11] Step 1 of Golden Gate assembly is cloning the desired sgRNA (purple, red, pink, light pink boxes) into the backbone (pX330A_D10A1x4) or insert vector (pX330S_2-4). Each vector contains Cas9 (green rectangle), a sgRNA scaffold (maroon rectangle), a U6 promoter (blue arrow), and a resistance marker either Ampicillin (yellow rectangle) for backbone or Spectinomycin (orange rectangle). Each vector contains BbsI cut sites (red dash lines) that create unique overhangs for scar less cloning. Step 2 proceeds by BsaI digestion (light blue dashed lines) removing the U6 promoter, sgRNA, and sgRNA scaffold and ligating into the backbone vector creating one multiplexed vector. (B) Example of colony PCR following Step 1 of Golden Gate. A forward primer against the U6 promoter and a reverse primer against the antisense of the sgRNA inserted amplify a 104 bp band. All samples except #5 were positive for sgRNA insertion. (C) Example of colony PCR following Step 2 of Golden Gate. A forward primer against the U6 promoter and a reverse primer against the final antisense of the sgRNA inserted amplify to create a band at 500 bp increments, with the complete vector of four sgRNAs being 2000 bp. Samples marked with an * indicate successful integration of all sgRNAs (D) AflIII and KpnI test digest confirms insertion of sgRNAs in multiplex vectors. Each sgRNA insertion increases size 500 bp, and digestion can confirm presence of incomplete or complete insertion of four sgRNAs 1 (sample B).

A




**Appendix Figure A.2: Integration of fluorescent proteins using Golden Gate CRISPR.** (A) Transfection of human iPSCs with sgRNAs designed against *CLYBL* locus for insertion of Ngn1/2 plus iRFP. After repeated rounds of fluorescent selection, homogeneous colonies were created for purposes of later differentiation into neurons for future experiments.

**Appendix Table A.1 Plasmids and primers**

| Name | Target | Source | Sequence |
|---|---|---|---|
| CLYBL_sgRNA_F | CLYBL | | F: CACCGACTTCCTTCTATGTAAGATG |
| | | | R: AAACCATCTTACATAGAAGGAAGTC |
| CLYBL_sgRNA_R | CLYBL | | F: GACCGATATTTATGTTGGAAGGATG |
| | | | R: AAACCATCCTTCCAACATAAATATC |
| LC3_sgRNA_F | LC3 | | F: CACCGCACCAATCTCAGAGGTGTAT |
| | | | R: AAACATACACCTCTGAGATTGGTGC |
| LC3_sgRNA_R | LC3 | | F: CACCGATGCCTCCCAGGAGACGTTC |
| | | | R: AAACGAACGTCTCCTGGGAGGCATC |
| U6_F | U6 promoter | | ACTATCATATGCTTACCGTAAC |
| GoldenGate_Step 2_F[11] | 5’ of first sgRNA | | GCCTTTTGCTGGCCTTTTGCTC |

| | | | |
|---|---|---|---|
| GoldenGate_Step 2_R[11] | 3’ of last sgRNA | | CGGGCCATTTACCGTAAGTTATGTA ACG |
| pX330A_D10A_1x 4[11] | | Addgene #1000000 055 | |
| pX330S_2[11] | | Addgene #1000000 055 | |
| pX330S_3[11] | | Addgene #1000000 055 | |
| pX330S_4[11] | | Addgene #1000000 055 | |