An integrated proteomic approach for mapping the ALS-linked TDP-43 interactome

By Anjali Patel

A thesis submitted in partial fulfillment of the requirements for the
Master's degree in Neuroscience

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

DEDICATION

This thesis is dedicated to *all health care workers, scientists, and volunteers* for their efforts in combatting the COVID-19 pandemic. To this end, I would like to acknowledge the contributions of my late aunt, Rashmita Patel, BSc.N. Her strength and resilience will continue to inspire me and her dedication to her patients has helped me recognize the end goal of my own pursuits— *to help others*.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest and sincerest gratitude to my parents for their endless love and words of encouragement. I am forever indebted to them for providing me with the opportunities and experiences that have shaped my path, and their instillment in me of the value of education and hard work. Along these lines, I am also thankful to my best friend and sister, Hemali Patel, for inspiring me with her positive outlook on life and for her ability to make me laugh, in any circumstance. My family has been my greatest blessing and strongest pillars of support. I would not have reached this milestone or any others without them, and this thesis is a testament to their unyielding support of my goals and endeavours.

I would like to express my gratitude to my research supervisor and mentor, Dr. Maxime W. C. Rousseaux (University of Ottawa), for giving me the opportunity to conduct this research and for the guidance he has provided throughout my master's studies. His support in the lab when I faced experimental obstacles has taught me to persevere and his encouragement toward his students to become involved in extracurricular and STEM outreach initiatives has highlighted to me the importance of science communication and instilled in me the value of being an involved member of the community. The lessons that Dr. Rousseaux has helped me to learn expand beyond the realms of scientific research and are ones that I will carry forward in my life and career.

I would like to express my deepest gratitude to my co-supervisor, Dr. Johnny K. Ngsee (University of Ottawa) for his advice and support throughout my studies, and the boundless enthusiasm for my project that he expressed throughout our discussions. In addition to these professors, I would like to thank the members of my thesis advisory committee, Dr Derrick Gibbings (University of Ottawa) and Dr. Laura Trinkle-Mulcahy (University of Ottawa) for their helpful feedback, questions and insight which have helped guide my work.

I would like to thank my lab mentor, Terry R. Suk (PhD candidate, University of Ottawa) for his efforts in creating and characterizing the knock-in cell lines, without which this project could not have been

completed. I would also like to express my deepest gratitude for the time he spent teaching molecular lab techniques, providing feedback for optimization, troubleshooting and data analysis, and in providing his support throughout the duration of my studies.

I would like to express my heartfelt gratitude to my colleagues, whom I consider my friends, in the Rousseaux lab, who have all been an unyielding source of encouragement and support through the trials and tribulations of graduate school. I would also like to highlight the contributions of Rousseaux lab manager Steve M. Callaghan (University of Ottawa), who with his fountain of knowledge about scientific reagents & techniques and the workings of the University of Ottawa - Faculty of Medicine, has helped me navigate many situations. I would also like to express my thanks to my colleague, Haley M. Geertsma (PhD candidate, University of Ottawa) whose technical assistance with mouse dissections and words of encouragement have helped me persevere through my degree.

I would like to acknowledge the contributions of my collaborators, Emily Hashimoto-Roth (MSc candidate, University of Ottawa) for her time and assistance with data analysis for my project. Likewise, I would like to thank her supervisor, Dr. Mathieu Lavallée-Adam (University of Ottawa), for his feedback which helped guide the data analysis. I would like to acknowledge the work of Dr. Lawrence Puente (Ottawa Hospital Research Institute Proteomics Core, Ottawa) for providing workflow suggestions and for performing the mass spectrometry experiments. I would also like to thank Dr. Hilal Lashuel (Swiss Federal Institute of Technology in Lausanne) for his efforts in testing the α-synuclein variants for their aggregation potential.

Lastly, I would like to thank my friends--old and new--who have and continue to celebrate my accomplishments as if they were their own. It is with their kind words of support and wisdom that I am able to overcome the trials of life to pursue my goals.

FUNDING

I would like to express my sincerest gratitude to my funding sources, without whom this research would not have been possible. Firstly, I would like to acknowledge the **Eric Poulin Centre for Neuromuscular Disease**. As a "STaR award for ALS research" recipient, in addition to funding, I have had the opportunity to share my work with others during the University of Ottawa Brain Health Awareness Week.

In addition, I would also like to express my gratitude to the **Parkinson's Research Consortium** and the **Toth Family**. Through their generous funding of the "Toth Family Fellowship", I have had the funding to progress my research, as well as the opportunity to share my work with the Parkinson's research community through departmental seminars.

TABLE OF CONTENTS **PAGE NO.**

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder in which an RNA-binding protein, TDP-43, mislocalizes and pathologically accumulates from its normal nuclear locale to the cytosol[1,2]. Given that the subcellular localization and expression of TDP-43 is tightly regulated[3], we posit that identifying novel interactors of wild-type and mutant TDP-43 could reveal insight into networks involved in regulating its localization, ultimately driving neurodegeneration in ALS.

Using CRISPR/Cas9, our lab previously generated knock-in cell lines expressing GFP in the endogenous *TARDBP* locus (encoding for TDP-43) for both *wildtype* (WT) and an ALS-causing mutant (Q331K)[4]. We have shown that the Q331K mutation causes loss-of-function and mislocalization of TDP-43. I performed immunoprecipitation coupled to mass spectrometry (IP-MS) on this cell model to elucidate interactors of WT- and Q331K- TDP-43. Our data show that there is an overall loss of TDP-43 interactors in cells with the TDP-43^{Q331K} mutation. By setting statistical cut-offs for significance, we identified 34 shared and 12 unique interactors of TDP-43^{WT}. We used bioinformatic approaches to identify enriched pathways and literature searches to look for interactors relevant to TDP-43 and ALS pathobiology. Our shortlist of 14 candidates for validation included proteins involved in the nuclear mRNA export pathways, RNA binding proteins and proteins identified in other interactome studies and TDP-43 based screens.

Using orthogonal approaches, we show evidence of robust interaction of four top hits (PABPC1, HNRNPC, DDX39b and ELAVL1) with TDP-43^{WT}, and a significant decrease in the degree of interaction of HNRNPC, DDX39b and ELAVL1 with TDP-43^{Q331K}. Importantly, this decrease in interaction was only observed at the endogenous level, highlighting the importance of maintaining the steady state levels of TDP-43 in the cell for these assays. We characterized the effects of knockdown and overexpression of these four hits using protein-specific overexpression constructs and shRNAs and observed a significant increase in TDP-43 nuclear localization upon knockdown of these four hits, suggesting that there is a functional effect associated with hit knockdown. Overexpression or knockdown of the top hits in a splicing assay did not identify significant changes in TDP-43's splicing or RNA binding abilities, suggesting that these hits do not affect splicing function in our hit characterization studies.

Using this novel experimental tool and unbiased screen, we identified alterations in TDP-43 protein-protein interactions in the context of ALS and have generated tools to characterize their roles in cellular functions using knockdown and overexpression approaches. Together with the knock-in cells, these tools will allow us to gain insight into pathways involved in driving neurodegeneration, in the context of ALS.

ABBREVIATIONS

ADD2	B-adducin [gene]
ALS	Amyotrophic lateral sclerosis
ALYREF	Aly/REF export factor [gene]
ANG	angiogenin [gene]
ATXN2	ataxin-2 [gene]
Bbc3	BCL2 binding component 3/Puma [gene]
Bax	BCL2 Associated X, apoptosis regulator [gene]
C9ORF72	chromosome 9 open reading frame 72 [gene]
CHTOP	chromatin target of PRMT1 [gene]
CSF	cerebrospinal fluid
CTD	c-terminal domain
CTF25	c-terminal fragment of 25 kDa
CTF35	c-terminal fragment of 35 kDa
DDX39b	DExD-Box Helicase 39B [gene]
ELAVL1	elav-like RNA binding protein 1 [gene]
EMF	electromagnetic field
EMG	Electromyography
fALS	familial ALS
FDR	false discovery rate
FRG1	FSDH region gene 1 [gene]
FTD	frontotemporal dementia
FUS	Fused in Sarcoma [gene]
GFP	Green fluorescent protein [gene, protein]
GFP-IP	GFP-immunoprecipitation
HCA	hierarchical clustering analysis
HDR	Homology directed repair
HNRNPA0	heteronuclear ribonucleoprotein A0 [gene]
HNRNPA2B1	heteronuclear ribonucleoprotein A2B1 [gene]
HNRNPC	heteronuclear ribonucleoprotein C [gene]
ILF3	interleukin enhancing binding factor 3 [gene]
IPA	Ingenuity Pathway Analysis software
IP-WB	Immunoprecipitation-western blot
KD	Knockdown
KI	Knock-in
LCD	Low complexity domain
LC-MS/MS	liquid chromatography tandem mass spectrometry
MATR3	matrin-3 [gene]
MG	Myasthenia Gravis
MRI	magnetic resonance imaging
NES	nuclear export signal
NLS	nuclear localization signal
NME2	NME/NM23 Nucleoside Diphosphate Kinase 2 [gene]

NTD	n-terminal domain
OE	Overexpression
OPTN	optoneurin [gene]
PABPC1	Poly(A) Binding Protein Cytoplasmic 1 [gene]
PBA	Pseudobulbar affect
PCA	principal component analysis
PLS	Progressive Lateral Sclerosis
POLDIP3	Polymerase delta interacting protein 3 [gene]
PPP1Ca	Protein Phosphatase 1 Catalytic Subunit Alpha
RRM1	RNA recognition motif 1
RRM2	RNA recognition motif 2
SAINT	Significance Analysis of INTeractome Program
sALS	sporadic ALS
SARNP	SAP Domain Containing Ribonucleoprotein [gene]
sfGFP	Superfolder green fluorescent protein [gene]
SINE	selective inhibitors of nuclear export
SMA	Spinal Muscular Atrophy
SNRNP200	Small Nuclear Ribonucleoprotein U5 Subunit 200 [gene]
SOD1	Zn/Cn Superoxide dismutase
SQSTM1	Sequestosome 1 (also known as ubiquitin binding protein p62)
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins Program
ssODN	Single stranded deoxyoligonucleotide template
TARDBP	transactive response element DNA-binding protein of 43 kDa [gene]
TIA1	Tia1 cytotoxic granule-associated RNA binding protein [gene]
TDP-43	protein, encoded by TARDBP gene
VAPB	Vesicle-associated membrane protein-associated protein B/C [gene]
WT	*wildtype*
XPO1	exportin 1 [gene]
YBX1	Y-Box Binding Protein 1 [gene]

LIST OF FIGURES AND ILLUSTRATIONS

SUPPLEMENTARY FIGURES

LIST OF TABLES **PAGE NO.**

1. AMYOTROPHIC LATERAL SCLEROSIS

1.1 HISTORY OF ALS

Amyotrophic lateral sclerosis (ALS) is a progressive, late-onset neurodegenerative disease that results in the loss of motor neurons that control voluntary movement[5]. Clinical features of ALS have been written in many accounts including those of British surgeon Charles Bell, who may have been the first to capture its neurological symptoms in 1824[6]. This was later followed by French physician, François-Amilcar Aran, who in 1850 described 'progressive muscular atrophy', a disease whose symptoms matched the clinical features of ALS[7]. The disease was first termed as '*amyotrophic lateral sclerosis'* in 1869 by French neurologist Jean-Martin Charcot, recognized as 'the founder of modern neurology', who used his detailed clinical records to correlate the symptoms seen in his patients to the degeneration found in the anterior horn of the spinal cord during post-mortem tissue analysis[5,8]. In more recent years, the diagnosis of American professional baseball player, Lou Gehrig, shed public attention on this disease and resulted in it being also recognized as 'Lou Gehrig's disease' colloquially[5,8].

1.2 DIAGNOSIS AND PRESENTATION OF DISEASE

1.2.1 DIAGNOSIS

Patients are diagnosed with ALS through taking a clinical history, regular neurological assessments to assess the progression of muscle weakness and wasting and ancillary tests to rule out other diseases that share symptoms of ALS[9]. Some of the tests used in the diagnosis of ALS include electromyography (EMG), to detect electrical activity of muscle fibers, muscle biopsies to rule out other muscular diseases, magnetic resonance imaging (MRI), to image the brain and spinal cord, and blood and urine tests to rule out other diseases, among others[9,10]. Given that ALS cases are predominantly sporadic (sALS) with no known family history, genetic testing is rarely used for the diagnosis of this disease[11]. However, in individuals with known family history of ALS or that were diagnosed at a younger age, genetic tests can be used to identify the presence of known ALS-linked mutations in several genes[12].

1.2.2 NEUROLOGICAL SYMPTOMS

In ALS, the death of motor neurons in the central nervous system results in symptoms affecting movement. The onset of ALS can occur in two distinct ways—limb onset ALS or bulbar onset ALS[12]. In limb onset ALS, symptoms such as muscle stiffness, twitching, muscle atrophy, and loss of grip strength initially occur at the arms and/or legs. Early signs of developing these symptoms include loss of balance or difficulties performing manual dexterity actions[12,13]. Limb onset ALS accounts for most ALS cases, and over time, patients progress to develop symptoms affecting the face, neck and respiratory system. Bulbar onset ALS initially occurs in the face or neck, and results in symptoms such as slurred speech, difficulties chewing and swallowing, and weakness of facial, throat and jaw muscles[13,14].

ALS symptoms may initiate at one site of the body (focal) or result in symptoms initiating at multiple sites (multi-focal)[13]. The presentation of ALS symptoms differs between individuals and progression of disease can vary; however, bulbar onset ALS typically progresses faster[14]. At end-stages (on average ~2-5 years after diagnosis), patients typically die of respiratory failure and pneumonia[5]. In addition to these physical symptoms, around 15-30% of patients can also experience cognitive and behavioural impairments[5].

The symptoms of ALS are similar to other neurological diseases affecting movement including frontal temporal dementia (FTD), motor neuron disease (NMD), progressive lateral sclerosis (PLS), spinal muscular atrophy (SMA), Myasthenia gravis (MG), among others[15-16,17]. In addition to the overlap in neurological symptoms between ALS and other diseases, there is also shared clinical pathology such as the presence of ubiquitinated TDP-43-protein-positive cytoplasmic inclusions such as those seen in FTD, which can make it hard to diagnose individuals with ALS[18].

1.2.3 CLINICAL PATHOLOGY

A clinical hallmark of ALS is the formation and presence of cytosolic protein aggregates which have been linked to neurodegenerative processes leading to motor neuron death[1]. There is a great deal of heterogeneity seen in ALS patients with regards to the underlying cause and disease presentation; however, a commonality seen in over 97% of ALS patients, is the formation of these cytosolic protein inclusions primarily comprised of the DNA/RNA binding protein, TDP-43[1,7-8].

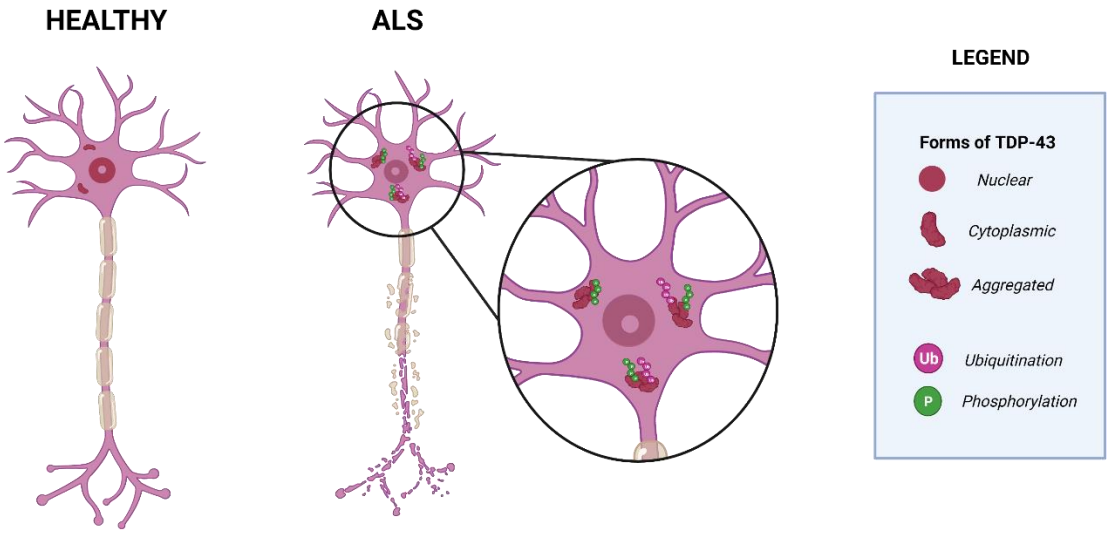


Figure 1: TDP-43 localization in healthy vs. ALS motor neurons. Graphical schematic of motor neurons from healthy [left] vs. ALS [right] individuals. TDP-43 is primarily nuclear in healthy cells. Hyperphosphorylated and polyubiquitinated TDP-43-positive cytosolic aggregates are seen in cells of ALS patients.

The mixture of full-length and truncated TDP-43 within these aggregates is usually found to be hyperphosphorylated and polyubiquitinated (figure 1). A small subset of ALS patients (<3%), including those with mutations in the *SOD1* (Zn/Cn Superoxide dismutase) and *FUS* (fused in sarcoma) genes, do not present with cytosolic TDP-43 inclusions[1,5,8].

1.3 EPIDEMIOLOGY

The global incidence rate of ALS is around 1.5 in 100,000 and this incidence rate has remained relatively constant over the last decade[20]. In Canada, there are approximately 3,000 adults currently living with ALS[21]. Studies have shown that males have an increased risk of disease (1.5:1, males to females)[22]. ALS typically affects older individuals (average age of 50-65 years, with increased risk correlating with age), although a smaller proportion of cases (~5%) are early onset. ALS cases can be categorized as either familial (fALS) or sporadic (sALS)[5]. fALS cases account for less than 5% of patients, while most cases are sporadic--with no known family history[5].

Epidemiological and GWAS studies have identified several risk factors, encompassing both genetic causes and environmental factors that are associated with the disease[23].

1.3.1 GENETIC CAUSES

Among the genes identified in ALS GWAS and exome sequencing studies, the most well known and studied include *C9ORF72 (Chromosome 9 open reading frame 72)*, *SOD1*, *TARDBP* (*Transactive response element DNA-binding protein of 43 kDa*) and *FUS. C9ORF72* mutations, caused by hexanucleotide repeats in the gene, account for most ALS cases (40% of fALS and 7% of sALS)[24,25]. A mutation in the *SOD1* gene was the first to be discovered, and this form of ALS clinically presents with the formation of SOD1-containing protein aggregates, lacking the ubiquitinated TDP-43 protein aggregates seen in the majority of ALS patients. Mutations in the *TARDBP* (encoding for the TDP-43 protein) and *FUS* genes account for about 4% of fALS cases each, and both gene mutations account for < 1% of sALS cases altogether. Some of the other genes that have also been linked to ALS include *MATR3, ANG, OPTN, HNRNPA0* and *VAPB,* among others[26,27]. It has been suggested that mutations in these genes may result in impaired cellular stress pathways, perturbations in RNA machinery and processing, apoptosis, impaired axonal transport, protein aggregation, mitochondrial dysfunction, excitotoxicity, and impaired glutamate handling[28].

With the general exception of *SOD1* and *FUS* mutation carriers, mutations in these ALS-linked genes culminate in TDP-43 pathology—making the study of TDP-43 pathobiology a growing aspect of ALS research. Aspects of TDP-43 pathology seen in ALS include the formation of characteristic hyperphosphorylated and polyubiquitinated TDP-43 cytoplasmic inclusions, which may also include truncated TDP-43 protein variants[29]. In a 2009 study by *Lee et al*, the phosphorylation of TDP-43's S409/410 residues were identified to be pathological hallmarks of lesions seen in TDP-43 proteinopathies such as ALS[30]. These inclusions have also been found to colocalize with markers of autophagy and the protein, p62/SQSTM[31,32]. In addition to these cytoplasmic aggregates, other avenues of TDP-43 pathology include its nuclear depletion and cytoplasmic mislocalization.

1.3.2 ENVIRONMENTAL FACTORS

ALS shows clinical heterogeneity in that patients with the same genetic mutations can have varied presentation in terms of age-of-onset, progression, symptoms, and pathology. This suggests that ALS is likely a result of interactions between an individual's genetic background and their surrounding environment. Studies have linked ALS to environmental risk factors including smoking, diet, physical activity, and exposure to chemicals/metals/electromagnetic fields[23].

1.3.2.1 Exposure to chemicals

Smoking

Several longitudinal studies have found that smokers have a higher risk of developing ALS. The exact cause underlying this increased risk is unknown but it has been suggested to be due to increased oxidative stress in motor neurons caused by nicotine or other toxins present in cigarettes and cigarette smoke[33].

Heavy metals

Several epidemiological studies have linked exposure to heavy metals such as lead, mercury, selenium and several others with an increased risk of ALS[34]. There are a large amount of data supporting the hypothesis that heavy metal exposure is neurotoxic. For example, studies have found populations of ALS patients in communities with high mercury exposure though diet and occupation[35,36]. Likewise, studies have found increased ALS risk in communities with tap water high in selenium[37]. Additionally, *in vivo* models have shown that exposure to high levels mercury or selenium can result in the onset of ALS-like phenotypes in transgenic mice, as well as motor neuron degeneration and motor function defects in swine and cattle[38-40]. Conflicting evidence suggests that lead exposure can prolong survival in $SOD1^{G39A}$ transgenic mice, which express an ALS-causing mutation and are used to study ALS phenotypes[41]. Conversely, longer survival was associated with lower lead levels in a more recent study, and there is a significant amount of literature that supports the idea that lead exposure is neurotoxic[42].

1.3.2.2 Diet

Studies in Japanese and Dutch cohorts have shown that diets high in vitamin E are associated with a decreased risk of ALS[43]. Other studies have also shown that longer use of vitamin E is associated with a

lower ALS risk. However, studies in smaller populations have shown no differences in serum or CSF levels of vitamin E between ALS patients and healthy individuals[44]. Vitamin E supplementation with riluzole treatment showed an improvement in slowing deterioration but did not improve survival in ALS patients[45]. In addition to vitamin E, other antioxidants such as vitamin C, carotenes, curcumin and coenzyme Q10 have shown some effect in improving ALS phenotypes in mouse models and human patients[46]. One such example is the drug edaravone (Radicava®), which has been FDA-approved for the treatment of ALS, and functions by reducing the effects of oxidative stress on cells by acting as a free radical scavenger of lipid and hydroxyl radicals[47,48].

1.3.2.3 Physical Activity

Evidence suggests an increased risk of ALS in individuals who participate in high intensity physical activity[49]. Several studies have shown an increased risk among professional, but not recreational football amd soccer players[50,51]. Although the exact underlying mechanisms are unknown, possible explanations for this increased risk include increased oxidative stress and glutamate excitotoxicity, both of which are suggested to result in the increased production of reactive oxygen species (ROS) and calcium influx—leading to cell death[52]. Potential factors for this increased risk include repeated head injuries, chemicals used to treat sports fields and vigorous physical activity. Clinical data also shows that ALS patients have lower BMI and higher levels of physical fitness compared to healthy individuals[53]. The evidence supporting the idea that high intensity physical activity is associated with ALS risk is inconsistent and these studies have been only conducted in smaller populations. Other studies have suggested that there is no link between physical activity and ALS risk, and few studies have suggested that physical activity could be neuroprotective[23,54].

1.3.2.4 Electromagnetic fields

There is also conflicting evidence suggesting that residential or occupational exposure to electromagnetic fields (EMFs) may lead to an increased risk of developing ALS[55]. Studies of European cohorts have found that electrical and engineering workers, who are continuously exposed to electromagnetic and electrical fields show an increased risk for ALS[55,56]. Data countering this hypothesis comes from larger population

studies done in the Netherlands which found that individuals living near high-voltage power lines (EMF sources) did not show an increased risk for ALS[57]. As well, more recent data has suggested that while an association between ALS and electrical occupations exists, this association cannot be explained by exposure to electrical or magnetic fields[58].

1.3.2.5 Limitations

Many of these studies have been conducted in smaller populations and present conflicting evidence on the contribution of the aforementioned risk factors in increasing the risk for ALS. Given that most of these correlational studies have been conducted in small populations residing in a given geographical area, it is important to recognize the effects of sampling biases and other confounding factors which may lead to inconsistences in the conclusions made between studies.

1.4 ALS MANAGEMENT AND TREATMENT

There is currently no cure for ALS, and disease management typically involves the management of pain and symptoms—with the intent of maintaining quality of life and prolonging lifespan[5]. Management strategies include multidisciplinary care, ventilation support, dietary changes, and analgesic and anxiolytic medications, among others[59].

There are a few FDA-approved drugs such as edaravone (Radicava), riluzole (Rilutek, Tiglutik, Exservan) and dextromethorphan/quinidine (Nuedexta) which are used for the treatment of ALS symptoms[60,61]. Edaravone, administered intravenously, has been shown to delay the decline of physical function in ALS patients[63]. Although the exact mechanism of action is not well understood, it is thought that edaravone acts as a free radical scavenger to relieve cellular oxidative stress, a pathological event that can lead to cell death[63]. Riluzole, administered orally, has been found to prolong survival (by three months on average) and delay the onset of ventilator support requirements in some patients[64,65]. It is thought that riluzole acts through preventing the release of the neurotransmitter glutamate, which has been shown to pathologically accumulate in the brain and spinal cord of ALS patients[66]. Dextromethorphan/quinidine is the only FDA-approved drug used for the treatment of pseudobulbar affect (PBA), a condition in which patients display frequent and involuntary outbursts of emotions that are exaggerated or incongruent with an emotional

stimulus[66]. Data suggests that PBA is exhibited in up to 50% of ALS patients[66,67]. In addition to these drugs, there are current ongoing clinical trails of many other drugs and clinical approaches to target ALS symptoms.

There is a limited understanding of the pathological mechanisms underlying ALS and current therapies are targeted at managing symptoms, at timepoints in the disease where there is a significant amount of neurodegeneration and loss of physical ability. Furthermore, treatment of symptoms does not significantly slow or halt the progression of disease. For these reasons, there is an unmet need for research focused on understanding the etiology and pathogenesis of ALS in order to develop new therapeutics that target the underlying disease-driving mechanisms.

1.5 TDP-43 PROTEINOPATHIES

TDP-43 neuropathology in the form of cytoplasmic, poly-ubiquitinated inclusions and neurites has also been observed in subsets of patients with frontotemporal dementia (FTD), which is a neurodegenerative disorder that is caused by progressive nerve cell loss in the brain's frontal and temporal lobes[68]. This causes a form of dementia characterized by changes in behaviour, speech and motor symptoms. FTD can co-occur with motor neuron or Lewy body diseases, and like ALS, the TDP-43 present in cytoplasmic inclusions is a mixture of hyperphosphorylated, polyubiquitinated and cleaved forms[69,70]. Other diseases that present with TDP-43-positive inclusions include limbic predominant age-related TDP-43 encephalopathy, Alzheimer's disease, and Lewy body dementias, among several others. Together, these disorders make up a subset of diseases classified as TDP-43 proteinopathies, and the findings from the present work may shed insight into perturbed pathways that are shared among these disorders.

2. TDP-43, AN ALS-LINKED RNA-BINDING PROTEIN

2.1 STRUCTURE AND FUNCTION

TDP-43 is a predominantly nuclear DNA/RNA-binding protein that can bind diverse RNA transcripts and has a broad range of functions including the regulation of mRNA splicing, RNA processing, stability, transport, and transcriptional and translational regulation[1,71]. TDP-43 is also found – at lower levels – in the cytosol where it has been suggested to be involved in RNA binding, translation and in regulating mRNA stability, autophagy processes and mitochondrial function[1,71,72]. It has also been shown to associate with components of stress granules in response to stressors[73].

The expression of TDP-43 is tightly controlled as the protein is capable of autoregulation through binding of the 3' untranslated region (UTR) of its own transcript, leading to the inclusion of an alterative intron to promote its degradation via nonsense mediated decay[74,75]. Other potential mechanisms that have been associated with TDP-43's binding to its own transcript as a means of autoregulation include increased mRNA instability, exosome mediated degradation and RNA polymerase II stalling or termination defects, among others[76]. This makes its study using overexpression or knockdown approaches particularly labour intensive and difficult to interpret. In addition to this, another approach through which physiological levels of this protein are maintained is suggested to be through its enzymatic cleavage into smaller fragments by cysteine proteases. Both C-terminal fragments of 25 (CTF25) and 35 (CTF35) kDa have been observed to be aggregation-prone in cells and have been shown to be a pathological signature of ALS brains[78-80].

TDP-43's structure can be organized into domains including the structured N-terminal domain (NTD), two highly conserved RNA recognition motifs (RRM1 and RRM2) and an unstructured, glycine-rich C-terminal domain (CTD)[1,81]. The protein's NTD has been shown to be responsible for its ability to reversibly oligomerize and is necessary for its splicing function. TDP-43 also possesses a nuclear localization signal (NLS), which is located between its NTD and RNA binding domains[1,82]. The presence of the NLS mediates its nuclear translocation, via active transport by the protein, importin α. Up until 2018, TDP-43 was also thought to have a nuclear export signal (NES), located within one of its structured RNA recognition motifs (RRM2), and it was thought that its mechanism of nuclear export was through binding of the protein,

exportin-1 (XPO1)[83]. However, more recent studies have found that inhibition of XPO1 using selective inhibitors of nuclear export (SINE) compounds or overexpression of XPO1 do not alter TDP-43 cytoplasmic localization[84]. Furthermore, the purported NES sequence is found within a structured domain of the protein, whereas typical NES sequences are in unstructured regions allowing for more effective XPO1 binding, which is not the case with the TDP-43 NES[82]. TDP-43's highly conserved RNA recognition motifs, RRM1 and RRM2, are responsible for its nucleic acid binding[71]. TDP-43's CTD, also known as the low complexity domain (LCD) has been shown *in silico* to be intrinsically disordered, and this prion-like region has the capability of forming beta-sheet structures which are characteristic of pathogenic aggregation[83,84]. The CTD is also required for TDP-43's splicing function and autoregulation abilities. This region also serves as the site of nearly all the protein's ALS-linked mutations, including the Q331K mutation, which is highlighted in this study[1,71]. In addition to this, many of TDP-43's protein binding partners including hnRNPs, UBQLN2 and FMRP, among others have been shown to bind on the CTD[82].

2.2 TDP-43 AND ITS IMPLICATIONS IN ALS

2.2.1 TARDBP ALS-LINKED MUTATIONS

3% of fALS and 2% of sALS cases can be attributed to mutations in the *TARDBP* gene, which encodes for the TDP-43 protein[85,86]. There are over 50 *TARDBP* mutations that have been identified, and most of them affect residues in TDP-43's CTD[87,88]. Studies have shown that the presence of these mutations can alter TDP-43's stability, propensity to aggregate, cytoplasmic localization, resistance to proteases and binding partners[82]. These mutations have varied prevalence within the population with A382T and M337V being the most abundant. Q331K, A315T, Q343R and M337V make up some of the most well-studied mutations and have been modelled using *in vitro* and *in vivo* models. These mutations have been found to increase the aggregation propensity and cytotoxicity of recombinant TDP-43 in yeast cells, and the A315T and M337V have also been found to increase cytoplasmic mislocalization[89,90].

The focus of this thesis, the TDP-43^{Q331K} mutation, was originally identified by Sreedharan *et al*. in a study examining TDP-43 mutations in British sALS patients[91]. The mutation was first identified in a 72-year old male with limb-onset ALS, and since then, has been modeled *in vitro* and *in vivo*. Compared to other

TARDBP mutations, it has been found to show robust phenotypes in cells, to be one of the most cytotoxic mutations when expressed in yeast cells, and its overexpression in mice has been found to result in non-lethal motor phenotypes, TDP-43 cytoplasmic aggregation and neuroinflammation[91,92]. Other studies have found that expressing this mutation in mice results in defects in motor function and motor neuron loss, as well as defects in synaptic transmission[93].

2.2.2 TDP-43 AGGREGATES

Cytoplasmic TDP-43-containing aggregates are seen in over 97% of ALS patients, irrespective of genetic background[1]. While these cytosolic aggregates are a histopathological hallmark of ALS, there is debate regarding whether these inclusions are a key driver of disease or a by-product of other dysfunctional cellular mechanisms. While the presence of aggregates is indicative of pathological mechanisms occurring within the cell and has generally been found to be toxic, there is also evidence against this idea. For instance, an *in vitro* study using a TDP-43 overexpression model in HeLa cells found that clearance of cytoplasmic TDP-43 aggregates does not mitigate cell death or cytotoxicity suggesting that there exist other mechanisms contributing to cell death[94]. Additionally, some studies report the formation of aggregates and subsequent sequestration of autophagy protein, p62 to be a mechanism of clearance—suggesting that this physiological protein clearance system is dysregulated in ALS[95-98].

2.2.3 TDP-43: ALS-LINKED HYPOTHESES

Some studies in the literature support TDP-43 mislocalization, which likely occurs prior to aggregation, as an avenue of TDP-43 mediated toxicity. Its nuclear-to-cytoplasmic mislocalization is suggested to contribute to pathogenesis through both loss-of-function and toxic gain-of-function mechanisms[82].

Its nuclear depletion is suggested to result in a loss of its normal nuclear function and its cytosolic accumulation into TDP-43-containing inclusions is suggested to confer a toxic gain-of-function in its new cytosolic locale[99]. An *in vitro* overexpression study using a TDP-43-ΔNLS construct to restrict exogenous TDP-43 to the cytoplasm observed its accumulation and sequestration into cytoplasmic aggregates. Other papers also show that increased aggregation of TDP-43-ΔNLS results in its nuclear depletion, insolubility and co-aggregation with *wildtype* TDP-43[98-100]. Furthermore, a paper by *Walker et al.* found that expression

of humanized TDP-43-ΔNLS in mice led to the formation of insoluble, phosphorylated cytoplasmic aggregates with resulting brain atrophy, muscle denervation, motor neuron loss and progressive motor impairments leading to death[101]. These are all features of the pathology seen in ALS patients. This same study found that Doxycycline-mediated suppression of TDP-43-ΔNLS after onset of pathological phenotypes led to a decrease in phosphorylated TDP-43, an increase in nuclear TDP-43 to baseline levels and prevention of further motor neuron loss[101]. Taken together, these studies support the idea that TDP-43's nuclear depletion, resulting from impairments of its nuclear-to-cytoplasmic shuttling mechanisms, leads to formation of cytoplasmic aggregates and pathology reminiscent of ALS.

Additionally, studies in various cell and animal models have shown that TDP-43 loss-of-function, due to its deletion or mislocalization, results in pathological or lethal events[99]. For instance, *in vitro* studies have shown that knockdown of TDP-43 results in abnormal cell morphology and increased cell death. Furthermore, TDP-43 knockdown studies in drosophila resulted in impaired larval locomotor activity, axonal loss and altered synaptic boutons[102-104]. Additionally, TDP-43 conditional knockout mice exhibit progressive motor impairments, motor neuron atrophy and degeneration at neuromuscular junctions, similar to the pathological events occurring in ALS patients[105].

Finally, TDP-43 cytoplasmic accumulation may result in toxic gain-of-function effects through the sequestration of mRNA transcripts and cytoplasmic proteins, abnormal nuclear-to-cytoplasmic shuttling and local translation[106]. The increased cytoplasmic presence of TDP-43 may also result in an increase of its normal cytoplasmic functions—conferring a 'normal' gain-of-function. In a paper by *Hansen et al*, overexpression of TDP-43 in drosophila led to retinal neurodegeneration, motor dysfunction and reduced lifespan[107]. A paper by *White et al.* found that TDP-43's autoregulation is perturbed in the TDP-43^{Q331K} mouse model, resulting in a 'gain-of-function' of TDP-43 and altered splicing of a dementia-associated gene, *Mapt*[108]. These mice also presented with FTD-like phenotypes and an altered cortical transcriptome where multiple ALS-linked genes were dysregulated[108]. These findings suggest that ALS-causing *TARDBP* mutations may cause changes at the transcriptome level with implications for altering the cell's native proteome and downstream pathways.

2.3 TDP-43 INTERACTOME AND PROTEOMICS

Given that TDP-43 is central to ALS and is involved within many cellular processes, there have been several studies that have used mass spectrometry-based proteomics to better understand the protein's biology.

A study by *Freibaum et al*. that aimed to uncover the global *wildtype* TDP-43 interactome using an overexpression model found that the protein interacts with many nuclear proteins including hnRNPs, RNA helicases, splicing factors, RNA transport proteins, translation and transcription factors and cytoplasmic proteins including ribosomal and mitochondrial proteins[109]. Mapping these interactions using STRING showed the presence of two distinct protein interaction networks—a nuclear/splicing cluster and a cytoplasmic/translation cluster. Similarly, a paper by *Ling et al.* used an immunoprecipitation mass spectrometry approach to find that *wildtype* TDP-43 mostly complexes with other members of the hnRNP family, as well as snRNP proteins, among other RNA binding proteins[110]. Furthermore, in a paper by *Sephton et al*, endogenous TDP-43 was isolated from rat brain nuclear extracts and processed using mass spectrometry[111]. This led to the identification of 25 nuclear interactors of TDP-43, of which 9 were not previously reported, including *MECP2, PTBP2, CALML3, EIF5A* and *GM9242* (similar to *HNRNPA3*)[111]. Overall, these nuclear interactors are involved in RNA binding, splicing and translation providing further evidence for TDP-43's known nuclear RNA-binding functions[112]. A study by *Kawaguchi et al.* suggested a potential role for TDP-43 and FUS in the DNA damage response after observing its binding to components of DNA repair machinery upon the induction of DNA damage[113]. Taken together, these studies shed light into *wildtype* TDP-43's protein-protein interactions and provide support for studies completed to uncover TDP-43's native functions.

There have been very few attempts at mapping the ALS-linked TDP-43 interactome in current literature. In a 2020 study by *Feneberg et al*, using a mutant TDP-43$^{M337V/-}$ transgenic mouse model, the interactomes of TDP-43^{WT} and TDP-43^{M337V} were compared using mass spectrometry[114]. Alterations in the mutant TDP-43 interactome were found, specifically that the mutation impaired protein-protein interactions involved in stress granule formation and endosomal-extracellular transport and showed reduced binding of translation initiation factors[114].

Many of these proteomics studies were accomplished using overexpression models or by artificially altering the cellular stress response or inducing DNA damage. Given that TDP-43 protein levels are tightly regulated, overexpression models may result in interactome changes that do not reflect normal TDP-43 biology. Using a model in which TDP-43 is expressed at endogenous levels may result in better interactome mapping of native protein-protein interactions. Furthermore, existing interactome studies have aimed to map the protein-protein interactions of *wildtype* TDP-43 and, apart from the *Fenebe*rg *et al*. study, which was completed using a transgenic mouse model, there have not been any attempts (to our knowledge) to study how ALS-linked TDP-43 mutations alter its interactome. Therefore, our understanding of how ALS-linked mutations alter the TDP-43 interactome is still largely unknown, and the use of an endogenous TDP-43-expressing model may be best suited to address this question.

3. HYPOTHESIS

ALS-linked mutations, which are known to affect TDP-43's subcellular localization, may alter its interactome—resulting in aberrant pathways that drive neurodegeneration.

4. AIMS

1) Use an unbiased approach to map the *wild-type (*WT) and ALS-linked TDP-43 interactome.
2) Characterize the cellular mechanisms underlying proteins and pathways involved in TDP-43 mislocalization and toxicity.
    a. Perform an integrated bioinformatics analysis to delineate candidates for validation and characterization.
    b. Validate and characterize top hits in the context of ALS-relevant cellular phenotypes.

5. CONTRIBUTIONS TO PROJECT

Generation and characterization of knock-in (KI) cell lines, including all initial microscopy and qPCR experiments was completed by Terry Suk (PhD. Candidate, Rousseaux Lab, University of Ottawa). Protein identification using LC/MS-MS experiments including preparation of peptides from protein samples, mass spectrometry runs, and library matching was done by Dr. Lawrence Puente (Ottawa Hospital Research Institute Proteomics Core, Ottawa, ON). Writing of 'protein identification by LC-MS/MS' section in *Methods* was included as part of the service's technical report and included in thesis with permission. Statistical analysis of mass spectrometry data, including writing and running code for PCA, HCA and SAINT analyses was completed by Emily Hashimoto-Roth (MSc. Candidate, Lavallée-Adam Lab, University of Ottawa). Writing of 'mass spectrometry data analysis' section in *Methods* was completed by Emily Hashimoto-Roth, reviewed by Dr. Mathieu Lavallée-Adam (University of Ottawa), and included with permission. The completion of all other experiments, including their data analysis and figure generation was completed by myself.

6. METHODS

1. GENERATION OF KNOCK IN (KI) CELL LINES

Cas9/sgRNA plasmid cloning. sgRNAs were designed using Benchling for TDP-43^{WT}-GFP knock-in cell lines. The sgRNA specifically recognizes the stop codon in the 3' end of exon 6 of the *TARDBP* gene. 1 ug of LentiCRISPRv2 (LCV2) plasmid was digested with BsmBI restriction endonuclease (New England Biolabs) as per manufacturer's protocols. Digested plasmid was run on 1% agarose gel, gel purified and phosphorylated. Annealing of sgRNAs was completed in thermocycler with following protocol: 37º C for 30 mins. 95ºC for 5 mins. Ramp down to 25ºC by rate of -5ºC/min. Hold at 4ºC. sgRNA and LCV2 plasmid were ligated (NEB Quick Ligase) for 10 mins at room temperature (RT) as per manufacturer's protocol. Reaction was transformed into chemically competent *e. coli* DH5α cells, by adding reaction mixture to cells and incubating on ice for 20 mins. After 20 min incubation, cells were heat shocked at 42ºC for 45 secs, and left to recover on ice for 2 mins, before adding S.O.C (Thermofisher Scientific) media, and letting culture grow at 37ºC for 1 hour. After 1 hour, cells were spun down and plated on appropriate ampicillin-LB agar plates and left to grow for 16-20 hrs at 37ºC. Colonies were selected, grown in LB media at 37ºC for 16-20 hrs, miniprepped and sent for sequencing using H-U6 primer to the Ottawa Hospital Research Institute StemCore Sequencing Facility (Ottawa, Canada).

Transfection of Homology Directed Repair (HDR) insert. HDR insert was ordered as a 2 kb gene block (TARDBP-sfGFP) containing superfolder GFP[115] (sfGFP) and 600 bp homology arms to 3' UTR and 3' end of *TARDBP* gene in the pUC19 plasmid. pUC19 plasmid was digested using BamHI and HindIII and PCR clean up was performed. TARDBP-sfGFP was ligated into digested plasmid and transformed as per methods described above. Colonies were selected, miniprepped and sequenced to confirm presence of insert. Sequence confirmed plasmid was digested using HindIII and BamHI for use in lipofectamine 2000 (Thermofisher Scientific) transfection. Co-transfection was performed using HDR dsDNA insert and sgRNA/cas9 plasmid into HEK293T

cells, as per manufacturer's protocols. Cell were incubated for 1-3 days at 37C, 5% CO2 and then selected with puromycin, and recovered before limiting dilution was used to generate monoclonal colonies. Monoclonal colonies were grown in each well of 96 well plate using single celling technique, at 37C, 5% CO2 for 7-14 days and GFP+ cells were identified with microscope. GFP+ wells were grown to confluency in dishes and gDNA was isolated from these cells and quality control PCR was performed (using primers designed for cloning of HDR insert) and successful colonies were submitted for sequencing to confirm presence of the GFP knock-in. TDP-43^{WT}-GFP KI cells are heterozygous for the GFP KI.

TDP-43^{WT}-GFP cells were used to generate the TDP-43^{Q331K}-GFP knock cell line. Single stranded oligo deoxynucleotide template (ssODN) was designed to contain c.991C>A (p.Q331K) mutation and several synonymous mutations that cannot be recognized by the sgRNA after generation of KI cells. Similar transfection protocol was followed as above, using 10uM ssODN and differential hybridization PCR screening was completed using primers designed over the knock-in. PCR amplification the region of interest was done on both alleles to ensure presence of mutation, and it was confirmed that TDP-43^{Q331K}-GFP KI cells are homozygous for the Q331K mutation.

2. CHARACTERIZATION OF KI CELL LINES

a. MICROSCOPY

TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP cells were plated in 24 well dish on acid-washed glass coverslips (in 10% fetal bovine serum (FBS, Millipore Sigma) and antibiotic-antimycotic (ABAM, Thermofisher Scientific), coated in poly-d-lysine (50 ug/ml, incubated at 37ºC for 1 hr prior to plating. Coverslips were rinsed three times with sterile water and left to dry for four hours before cells were plated). Cells were incubated at 37ºC for 72 hrs. After 72 hr incubation, cells were fixed using 8% paraformaldehyde (PFA, final 4% dilution) on rotational shaker for 20 mins at room temperature. Coverslips were washed three times with PBS, and primary antibody (αTDP-43, diluted in blocking buffer) was applied (1:1,000, see table 10) overnight, and coverslips were covered and incubated at 4ºC. 24 hrs later, coverslips were washed three times with PBS, and

incubated away from light in secondary antibody (diluted in blocking buffer) on rotational shaker for 1 hr at room temperature. Coverslips were washed three times in PBS and mounted on slides using antifade mounting media containing DAPI (Vectashield). After 10 min incubation away from light, coverslips were sealed using clear polish, and stored at 4ºC in anticipation of imaging on the Zeiss M2 fluorescent microscope (Cell Biology and Image Acquisition Core, Ottawa, Canada.).

b. QPCR

qPCR primers were designed using the Roche Probe Library Design Algorithm (table 1, table 6). RNA was isolated from cells using TriZol RNA isolation protocol with TriZol reagent (Thermofisher Scientific). cDNA synthesis was completed using iSCRIPT cDNA Mastermix (Bio-Rad), as per manufacturer's protocol. The following PCR conditions were used: priming at 29ºC for 5 mins, reverse transcription at 46ºC for 20 mins. Inactivation of reverse transcription at 95ºC for 1 min, hold at 4ºC before use. qPCR samples were plated using 10 µM mix of forward and reverse primers, cDNA and GoTaq Green2Go MasterMix (Promega), as per manufacturer's protocols. Plate was spun down at 1,000xg and read using the thermocycler.

3. GFP-IMMUNOPRECIPITATION

Cells were grown in DMEM (Wisent Bioproducts) media with 10% FBS and ABAM. Fully confluent 150mm plates of cells were lysed in NETN (100 mM NaCl, 20 mM Tris-Cl (pH 8.0), 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40, protease inhibitor, phosphatase inhibitor) lysis buffer, and vortexed 4 times, over 20 mins. Lysates were centrifuged at 13,000 rpm for 20 mins and applied to 30 µl GFP-trap magnetic Sepharose beads (Bulldog Bio) for 1 hr at 4ºC. Beads were washed 3 times using 0.5X NETN buffer and eluted using 2X laemmli buffer using thermomixer (90ºC for 10 mins, 950 rpm).

4. MASS SPECTROMETRY SAMPLE PREPARATION

Samples were run on 4-12% gels (BioRad NuPage PreCast Protean minigels) and silver stained (ThermoScientific Pierce Silver Stain for Mass Spectrometry) as according to manufacturer's

protocols. Gel bands were excised and stored in 1% acetic acid before submission to the Ottawa Hospital Research Institute Proteomics Core (Ottawa, Canada).

5. PROTEIN IDENTIFICATION BY LC-MS/MS

Proteomics analysis was performed at the Ottawa Hospital Research Institute Proteomics Core Facility (Ottawa, Canada). Proteins were digested in-gel using trypsin (Promega) according to the method of Shevchenko[116]. Peptide extracts were concentrated by Vacufuge (Eppendorf). LC-MS/MS was performed using a Dionex Ultimate 3000 RLSC nano HPLC (Thermo Scientific) and Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). MASCOT software version 2.6.2 (Matrix Science, UK) was used to infer peptide and protein identities from the mass spectra. The observed spectra were matched against human sequences from SwissProt (version 2019-08) and against an in-house database of common contaminants. The results were exported to Scaffold (Proteome Software, USA) for further validation and viewing.

6. MASS SPECTROMETRY DATA ANALYSIS

a. MASS SPECTROMETRY DATA PRE-PROCESSING

Spectral count data was exported from Scaffold (version 4.8; Proteome Software, Inc., Portland, OR), filtering protein identifications from the purifications for known non-human experimental contaminants, human contaminants (e.g., keratin proteins), and decoy proteins. This resulted in the exporting of a mass spectrometry dataset that was identified and quantified using spectral counting a total of 676 proteins across the triplicate wild-type (TDP-43^{WT}), triplicate mutant (TDP-43^{Q331K}), and triplicate negative control (HEK293T) immunoprecipitation experiments. Spectral counts were normalized against the total spectral count for all proteins identified in an immunoprecipitation experiment.

b. UNSUPERVISED MACHINE LEARNING EVALUATION OF RELIABILITY OF EXPERIMENTS

To evaluate the behaviour of the immunoprecipitation experiments performed, two unsupervised machine learning algorithms were used. The first was a hierarchical clustering analysis (HCA)

using the complete linkage algorithm and the Euclidean distance as distance measure, to investigate the similarity between experiments. This initial assessment would give insight into whether the TDP-43^{WT} and TDP-43^{Q331K} immunoprecipitation experiments were like the negative controls. Therein, such experiments would be less reliable for subsequent analysis of the putative protein-protein interactions. The hierarchical clustering analysis was implemented in R (version 3.6.1), using the *pvclust* package[117] (version 2.2.0), with which a bootstrapping analysis was also performed to evaluate the robustness of the resulting clusters. To corroborate the results of this first clustering analysis, a two-dimensional principal component analysis (PCA) was performed. Similar to the hierarchical clustering analysis, experiments that are similar would cluster together, providing insight into the reliability of the immunoprecipitation experiments. The PCA was implemented in Python (version 3.7), using the *Scikit-learn* package (version 0.22.0).

c. IDENTIFICATION OF HIGH-CONFIDENCE PROTEIN-PROTEIN INTERACTIONS USING SAINT

The SAINT algorithm (Significance Analysis of INTeractome)[118] was used to assess the confidence of putative protein-protein interactions in the TDP-43^{WT} and TDP-43^{Q331K} immunoprecipitation experiments. To run the algorithm, the Swiss-Prot (accessed February 12, 2020) and TrEMBL (accessed February 14, 2020) UniProt databases[119] were locally downloaded to perform the necessary data pre-processing required for SAINT's execution, such as retrieving the protein sequence length of purified proteins. SAINT was executed using the *saint-spc-ctrl* program with recommended parameters *nburn* = 2,000 (number of burn-in period for the Gibbs sampling procedure training the algorithm), *niter* = 10,000 (number of iterations in Gibbs sampling), *lowMode* = 1 (minimizing the impact on the confidence assessment of high spectral count interactions), *minFold* = 1 (forcing separation of positive and negative distributions), *normalize* = 0 (without spectral count normalization). SAINT associates an FDR to each putative interaction to assess its confidence level. Dynamic thresholds were used to identify protein-protein interactions significantly and uniquely enriched in either the TDP-43^{WT} or TDP-43^{Q331K} immunoprecipitation

experiments. To be considered unique, a given protein-protein interaction must have been identified with a false discovery rate less than 15% in one experimental condition and more than 20% in the other. Such dynamic thresholding decreases the arbitrariness of identifying unique interactions, unlike using a single threshold for significance. Shared high-confidence protein-protein interactions were identified as those whose false discovery rates were less than 15% in both the TDP-43^{WT} and TDP-43^{Q331K} immunoprecipitation experiments.

7. GENERATION OF 3X-FLAG-GENE-OF-INTEREST CONTRUCTS

Entry vectors for gateway cloning (LR reaction) were selected from hORFEOME V8.1 entry and expression library[120] (Broad Institute) and site directed mutagenesis was used to add stop codons, where required (table 5, 6). LR recombination reactions (Invitrogen) were set up and performed as according to manufacturer's protocols and half-reaction was transformed into chemically competent *E. coli* TOP10 cells, by adding LR reaction mixture to cells and incubating on ice for 20 mins. After 20 min incubation, cells were heat shocked at 42ºC for 45 secs, and left to recover on ice for 2 mins, before adding S.O.C (Thermofisher Scientific) media, and letting culture grow at 37ºC for 1 hour. After 1 hour, cells were spun down and plated on appropriate antibiotic-LB agar plates and left to grow for 16-20 hrs at 37ºC. Colonies were selected, grown in LB media at 37ºC for 16-20 hrs, miniprepped and sent for sequencing to the Ottawa Hospital Research Institute StemCore Sequencing Facility (Ottawa, Canada).

8. ShRNA GENERATION

splashRNA software was used to generate 97-mer oligos for each gene target (table 7). PrimerBLAST was used to design forward (fwd) and reverse (rev) qPCR primers for each gene to be assessed (table 8). 97-mer oligos were PCR amplified using miR-Gib- fwd and rev primers. PCR samples with added dye were run on 2% agarose TAE gel at 100 V and product bands were gel purified using PCR gel extraction kit. pAAV-YFP-miR-E was digested using EcoRI and XhoI for 1 hr at 37ºC. Digested plasmid was run on 1% TAE gel and gel purified. Gibson assembly reaction was set up and run for 1 hr at 50ºC. Reaction was transformed into *E. coli* TOP10 bacteria using

protocol from section 7 (above). Colonies were selected, miniprepped and test digested using XhoI and HindIII on 2% agarose TAE gel. Colonies with dropout fragment were tested using qPCR and submitted for sequence confirmation.

9. VALIDATION OF MASS SPECTROMETRY HITS

a. CO-OVEREXPRESSION IP

Cells were grown in DMEM media with 10% FBS and ABAM. 500 ng 3xFlag-GOI plasmid was co-transfected with 500 ng pEGFP-TDP-43^{WT} or pEGFP-N2 plasmids in 100 mm dish of HEK293T WT cells at 50% confluency with lipofectamine 3000 (Invitrogen) as per manufacturer's protocols. Media change was completed 24 hrs post-transfection. 48 hrs post-transfection, cells were lysed in NETN lysis buffer, and vortexed 4 times, over 20 mins. Lysates were centrifuged at 13,000 rpm for 20 mins and applied to GFP-trap magnetic sepharose beads (Bulldog Bio) for 1 hr at 4ºC. Beads were washed 3 times using 0.5X NETN buffer and eluted using 2X Laemmli buffer using thermomixer (90ºC for 10 mins, 950 rpm).

b. GFP-IP IN KI CELLS

Cells were grown in DMEM media with 10% FBS and ABAM. 2500 ng of 3xFlag-GOI plasmid was transfected in 100 mm dish of HEK293T WT, TFP-43^{WT}GFP, TDP-43^{Q331K}-GFP cells at 50% confluency with lipofectamine reagent (Invitrogen) as per manufacturer's protocols. Media change was completed 24 hrs post-transfection. 48 hrs post-transfection, cells were lysed in NETN lysis buffer, and vortexed 4 times, over 20 mins. Lysates were centrifuged at 13,000 rpm for 20 mins and applied to GFP-trap magnetic sepharose beads (Bulldog Bio) for 1 hr at 4ºC. Beads were washed 3 times using 0.5X NETN buffer and eluted using 2X Laemmli buffer using thermomixer (90ºC for 10 mins, 950 rpm).

c. WESTERN BLOT

Samples were run on 10% SDS-PAGE gel for 45 mins, 120V. Western blot transfer was completed at 340 mA for 90 mins at 4ºC. Nitrocellulose membranes (Bio-Rad) were

blocked in 5% milk powder (w/v) in tris buffered saline with 0.5% tween-20 (TBS-T) for 10 mins and washed 5 times. Primary antibodies were applied to membranes (see antibody table) and incubated overnight at 4ºC. 24 hrs later, membranes were washed in 1X TBST and incubated in secondary HRP-conjugated antibody (1:10,000) for 1 hr at room temperature. Clarity western ECL substate (Bio-Rad) was applied before imaging (exposure times vary).

10. MICROSCOPY

HEK293T cells were plated in 24 well dish on acid-washed glass coverslips (in DMEM media supplemented with 10% FBS and ABAM), coated in poly-d-lysine (50 ug/ml, incubated at 37ºC for 16 hrs prior to plating. Coverslips (Thermo Fisher Scientific) were rinsed three times with sterile water and left to dry for four hours before cells were plated). Cells were transfected with 500 ng of shRNA-expressing constructs, 24 hours post-plating using lipofectamine reagent (Thermo Fisher Scientific), as per manufacturer's protocols. Media change using DMEM supplemented with 2% FBS and ABAM was completed 24 hrs post-transfection, and cells were incubated at 37ºC for 72 hrs. After 72 hr incubation, cells were fixed using 8% paraformaldehyde (final 4% dilution) on rotational shaker for 20 mins at room temperature. Coverslips were washed three times with PBS, and primary antibody (αTDP-43, diluted in blocking buffer) was applied (1:1,000. see table 10) overnight, and coverslips were covered and incubated at 4ºC. 24 hrs later, coverslips were washed three times with PBS, and incubated away from light in secondary antibody (diluted in blocking buffer) on rotational shaker for 1 hr at room temperature. Coverslips were washed three times in PBS and mounted on slides using antifade mounting media containing DAPI (Vectashield). After 10 min incubation away from light, coverslips were sealed using clear polish, and stored at 4ºC in anticipation of imaging on the Zeiss M2 fluorescent microscope (Cell Biology and Image Acquisition Core, Ottawa, Canada.). Images were taken using monochromatic camera using the following channels: 647nm, EGFP, 507 nm (YFP), 488 nm, and DAPI with exposure adjusted as required for each individual experiment.

11. CELL PROFILER

Cell profiler[121] pipeline was designed to extract channel info from image headers for each image taken on the Zeiss M2 fluorescent microscope. Nuclei were defined as primary objects using the DAPI channel input. Cell bodies were defined as secondary objects using the YFP channel. Cytoplasm was defined as a tertiary object by subtracting nuclei from cell body. 647 nm channel was used to measure αTDP-43 signal intensity in nucleus and cytoplasm and to calculate cytosol to nuclear (C/N) intensity ratio. Output of program was mean αTDP-43 intensity and C/N ratio for each cell counted.

12. SPLICING ASSAY

6 well dishes of HEK293T cells grown in DMEM supplemented with 10% FBS and ABAM were plated at 50% confluency and after incubation of 24 hours at 37ºC, transfected with 2.5 µg of 3x-flag-hit or shRNA constructs using lipofectamine, as per manufacturer's protocol. After 24-hour incubation at 37ºC, media change using DMEM supplemented with 5% FBS and ABAM. 72-hours post transfection, cells were pelleted, and RNA was isolated using TriZol RNA isolation protocol with TriZol reagent (Thermofisher Scientific), as per manufacturer's protocols. cDNA was synthesized using 5X Reverse Transcriptase Master Mix (BioBasic), using PCR protocol of 15 mins at 42ºC, 5 mins at 85ºC and hold at 4ºC before use, as per manufacturer's protocols. PCR reaction was prepared using 2X GoTaq Master mix (Promega), 25 ng of cDNA, and 10 µM of *POLDIP3-e3* forward and reverse primers (see primer table 9). Reaction was run in thermocycler using the following PCR settings: initial denaturing for 2 mins at 95ºC, 35 cycles: denature at 95ºC for 30s, anneal at 50ºC for 20s and extend at 72ºC for 30s. Final extension was performed at 72ºC for 5 mins, and samples were held at 4ºC, prior to use. PCR samples were run on 5% TAE DNA PAGE gel for 1 hr at 100V, using 100 bp DNA ladder (Froggabio). Gel was incubated in 1X TAE buffer with RedSafe for 1 hour at RT, on orbital shaker. After 1 hr incubation, gels were imaged on GelDoc using EtBr setting, optimizing exposure for intense bands. Bands were quantified using

rectangular volume tool in Image Lab software program, taking ratio of lower variant-2 band (~555bp) to upper band (variant 1).

7. RESULTS

7.1 Generating stable GFP-expressing knock-in cells lines expressing TDP-43^{WT} and ALS-linked TDP-43^{Q331K} mutation to use as a cellular tool to study TDP-43 pathology in ALS.

To benefit from the strong antigenicity and versatility of the GFP tag without altering native TDP-43 function, we generated HEK293T cell lines expressing GFP on the c-terminus of endogenous, *wildtype* TDP-43 (TDP-43^{WT}-GFP KI cells) and an ALS-linked mutant, TDP-43^{Q331K} (TDP-43^{Q331K}-GFP KI cells) using the LentiCRISPR-V2 plasmid, which expresses sgRNAs and cas9. This plasmid was co-transfected with a homology directed repair double stranded DNA template that had homology to regions 600 bp upstream and downstream of the 3' end of the *TARDBP* gene, encoding for the TDP-43 protein (figure 2A). Cells were selected with puromycin and recovered, with subsequent limiting dilution methods used to generate monoclonal colonies. These colonies were screened for GFP+ cells and sequenced to confirm successful knock-in.

The Q331K mutant cell line was generated subsequently, using a single stranded oligo donor nucleotide (ssODN) template containing the Q331K and several silent mutations. This ssODN was used to introduce the Q331K mutation into the TDP-43^{WT}-GFP KI cell line. The presence of the silent mutations allowed us to screen colonies via a differential hybridization PCR screening approach.

PCR amplification of both alleles of the gene and subsequent sequencing confirmed that the Q331K-mutation is homozygous, while the GFP KI is heterozygous in both the TDP-43^{WT}- and -^{Q331K} GFP KI cell lines. Successful GFP-knock-in and expression was confirmed by western blot, and qPCR (figure 2B and 2C), which showed significant increases in GFP-expression in the TDP-43^{WT}-GFP KI cells, when compared to HEK293T cells.

To assess whether the addition of the *GFP*-KI resulted in alterations of TDP-43 function, we used qPCR to examine several genes that show altered expression patterns upon TDP-43 loss-of-function (figure 2C). In a paper by *Contessi et al*, *ADD2*, a gene encoding for β-adducin, was previously shown to have its expression regulated by TDP-43, through the binding and stabilizing of its mRNA transcript[122]. The authors of that study showed chimeric and endogenous *ADD2* downregulation upon TDP-43 depletion in HEK293

cells, modelling a loss of TDP-43 function. In our cells, knock-in of the sfGFP did not significantly alter *ADD2* expression, compared to control HEK293T cells. Similarly, a study by *Elden et al.* identified *ATXN2* to be a modifier of TDP-43-mediated toxicity, where dose-dependant increases in its expression resulted in increased retinal degeneration and enhanced TDP-43 protein accumulation in drosophila overexpressing WT- or Q331K- TDP-43[123]. We do not see any significant changes to *ATXN2* expression levels suggesting that the *GFP*-KI likely does not affect TDP-43 mediated toxicity.

A paper by *Vogt et al.* showed that expression of *wildtype* or mutant TDP-43 induces p53 targets and proapoptotic gene expression, suggestive of an active stress response[124]. Their *in vitro* experiments using neural progenitors showed increased expression of the BH3-proapoptotic genes, *Bax* and *Bbc3* (Puma) upon expression of *wildtype* and mutant TDP-43[124]. In our GFP-KI cells, we observed reduced, but insignificant, expression of the pro-apoptotic pathway gene, *BAX*, which could be attributed to the limiting dilution technique used to select for monoclonal colonies, as it may have selected for colonies with even mild pro-survival phenotypes. Taken together, these results indicate that there are no significant changes to endogenous TDP-43 levels or function with the KI of *sfGFP*.

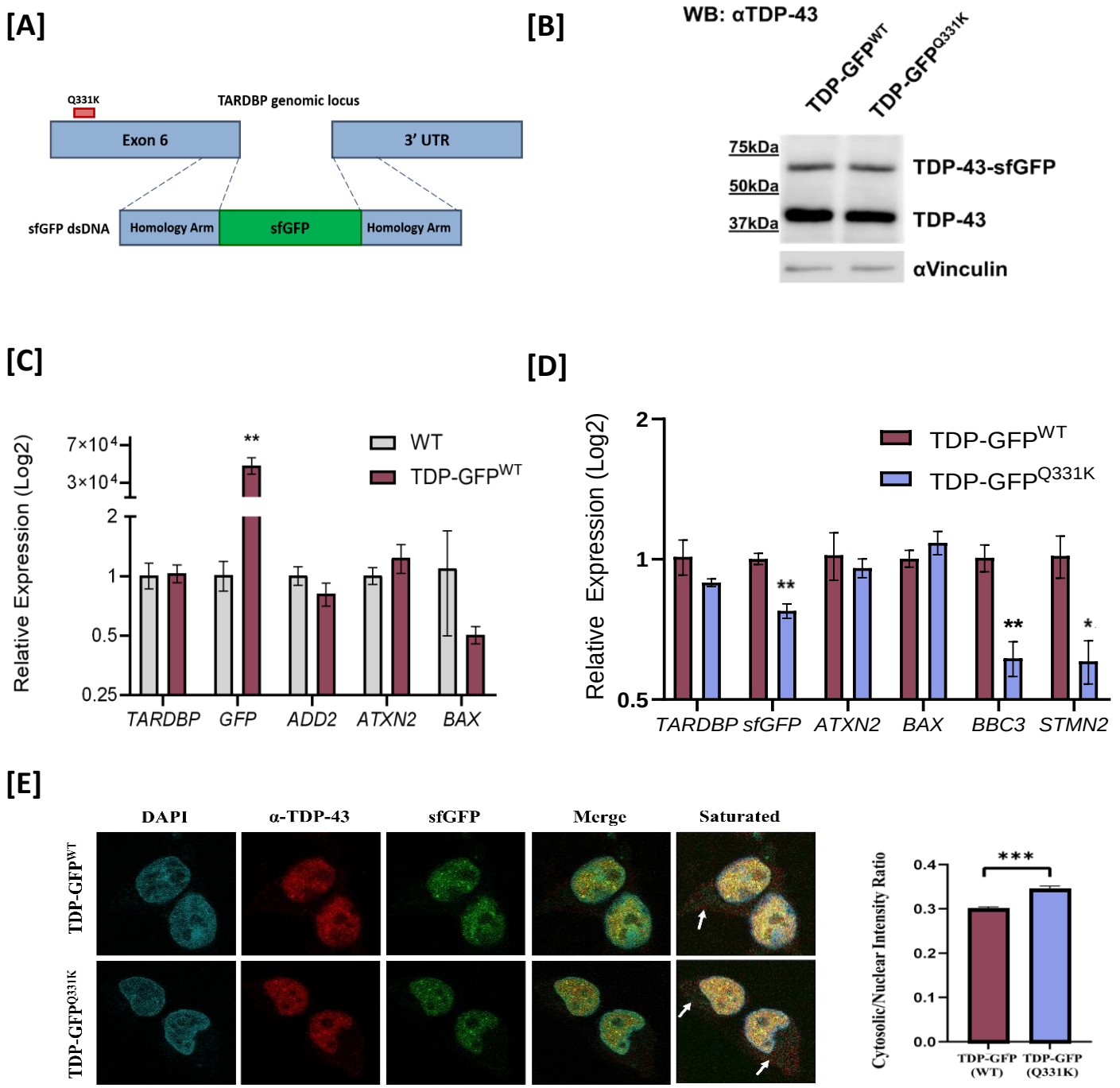


Figure 2 : Generation and characterization of TDP-43-GFP KI cell lines. [A] Schematic of Crispr/cas9 engineering approach used to generate TDP-43-GFP KI cell lines Red box represents Q331K mutation. **[B]** Western blot confirming equivalent protein-level TDP-43 expression in TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP KI cells. **[C]** Relative expression of select genes in HEK293T WT and TDP-43^{WT}-GFP cells, as analysed by qPCR, confirms increased GFP expression in the GFP-tagged cell line, with no significant effects on other genes **[D]** qPCR analysis of select genes in TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP KI shows downregulation of STMN2 mRNA, indicative of TDP-43 loss of function in the Q331K mutant cells. **[E]** [Left] Immunofluorescent staining of cell lines. TDP-43 is stained in red, DAPI in blue and GFP in green. The white arrows indicate TDP-43 localized in the cytoplasm. [Right] Graph plotting the cytosolic-to-nuclear TDP-43 ratio in both WT- and Q331K cells, suggesting a significant increase in cytosolic localization in the ALS-linked mutant.

Bars represent mean of experiments. Error bars indicate standard deviation. qPCR and microscopy experiments: (n=3), *p < 0.05, ** p < 0.01, *** p < 0.001 as determined by multiple unpaired t-tests with post-hoc analysis.

Experiments completed and figure generated by Terry R. Suk.

7.2 GFP knock-in cells lines expressing ALS-linked TDP-43^{Q331K} show TDP-43 loss-of-function and increased nuclear-to-cytoplasmic mislocalization.

To assess whether the introduction of the Q331K mutation alters TDP-43 function, we used qPCR to examine genes whose expression is regulated by TDP-43, or for whom changes in transcript levels are indicative of toxicity (figure 2D). There are no significant differences in *TARDBP, ATXN2* or *BAX* gene expression levels between the TDP-43^{WT}-GFP or TDP-43^{Q331K}-GFP KI cells. The expression of pro-apoptotic gene *BBC3* is significantly downregulated in the mutant cells, which may be explained by the use of the limiting dilution technique used to select for monoclonal colonies, as it may have led to the selection of cells with even mild pro-survival phenotypes, similar to the characterization of the TDP-43^{WT}-GFP KI cells, when compared to HEK293T cells.

sfGFP expression is significantly increased in the TDP-43^{WT}-GFP cells, however, this transcriptomic change does not translate over to the protein level (figure 2B and 2D). These results suggest that there are no significant differences in endogenous TDP-43 between the two cell lines, at the protein level.

In the TDP-43^{Q331K}-GFP KI cells, we observed a significant decrease in the expression of the *STMN2 gene,* encoding for the stathmin-2 protein (figure 2D). *STMN2* is a microtubule regulator discovered in an RNA-seq screen done by Klim *et al.* aiming to identify transcripts that are regulated by TDP-43[125]. In this study, expression of *STMN2* was reduced due to altered splicing after TDP-43 depletion, knockdown and mislocalization[125]. This decreased *STMN2* expression was also observed in spinal cord motor neurons taken from ALS patients[125]. Other studies have shown that TDP-43 regulates stathmin-2 mRNA by repressing premature polyadenylation, and that suppression of *STMN2* expression by premature polyadenylation is a hallmark of ALS[126,127]. In the *TDP-43^{Q331K}-GFP* cells, reduced *STMN2* expression indicates TDP-43 loss-of-function, which is reminiscent of the cellular-level pathology seen in ALS.

Using high content imaging microscopy, we have also observed the nuclear-to-cytoplasmic mislocalization of TDP-43 in the TDP-43^{Q331K}-GFP cells, quantified using Cell Profiler. Cell Profiler is an automated image analysis software which can be programmed to measure and quantify the intensity of the TDP-43 antibody signal in both the cytosol and nucleus of cells, in images. We observed a small, but statistically significant

increase in the TDP-43 cytosol-to-nuclear intensity ratio in the Q331K mutant cells (figure 2E) which is indicative of increased TDP-43 cytosolic localization, similar to the mislocalization seen in ALS.

7.3 GFP-immunoprecipitation coupled to mass spectrometry identifies groups of proteins unique to TDP-43^{WT}, showing overall loss-of-interactions associated with mutant TDP-43.

We tested and optimized a GFP-immunoprecipitation protocol using the knock-in cells to determine the best conditions for maximizing protein interactors, while minimizing non-specific binding to the Sepharose GFP-trap beads. To do this, we varied the number of cells used, the concentration and composition of lysis buffers and altered washing and elution conditions (refer to methods).

We performed GFP-immunoprecipitation using gentle lysis and washing conditions to identify proteins interacting with TDP-43^{WT} and TDP-43^{Q331K} using HEK293T cells as a background control for filtering out non-specifically binding proteins. We used three individual passages of cells to build our dataset, and these samples were run on SDSPAGE gels and silver stained to visualize protein pulldown. Gel slices were excised and submitted for mass spectrometry analysis (see methods for run details) to the OHRI Proteomics Core (Ottawa, ON). Peptides were matched to a spectral library of known proteins and common contaminant proteins such as keratin were removed from the dataset.

HCA analysis was performed between replicates of the experimental conditions, and it indicated that samples of the same conditions tend to cluster together, with a greater degree of clustering occurring between the TDP-43 IP samples than with control HEK293T cells (figure 3A). PCA analysis, which is used to measure the variance between datapoints showed that replicates of the same experimental condition tend to group together (figure 3B). Together, these findings indicate that each individual GFP-IP experiment has greater similarity to the others of the same experimental condition, and as such, it is unlikely that batch effects influence our dataset.

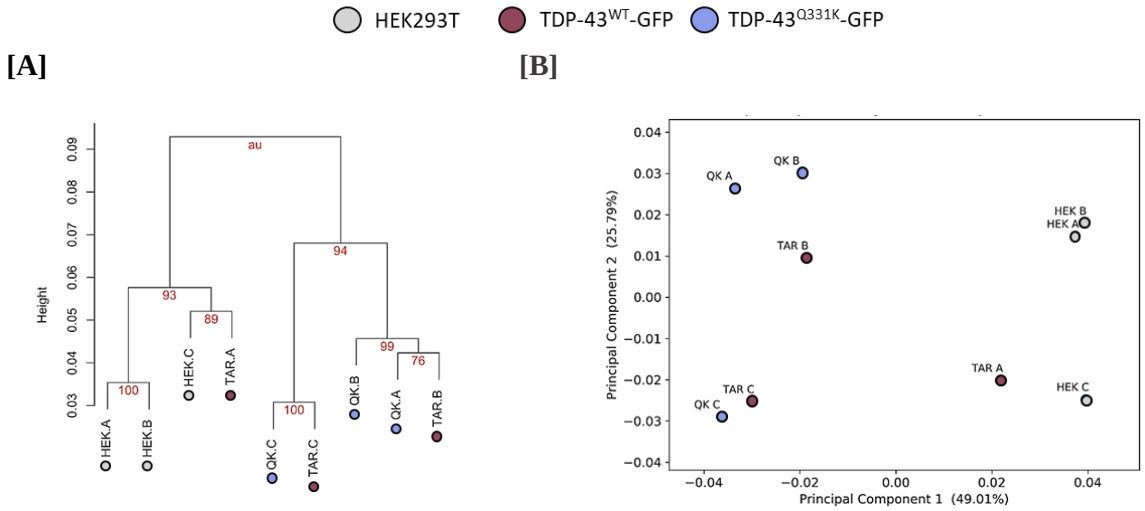


Figure 3: Assessment of data quality between replicates in each experimental condition indicates similarity and clustering between replicates of the same condition: [A] Hierarchical clustering analysis shows that replicates of the same experimental condition cluster together and are similar to each other. **[B]** Principal component analysis shows that there is less variance between replicates of the same condition than between replicate samples of other experimental conditions. Replicates have been coloured in grey (HEK293T), maroon (TDP-43^{WT}-GFP) and light purple (TDP-43^{Q331K}-GFP).

Data analysis for figure and figure generated by Emily Hashimoto-Roth and edited by Anjali Patel.

To determine significantly enriched proteins associated with each experimental condition (TDP-43^{WT} and TDP-43^{Q331K}), we used SAINT, which is a probabilistic proteomics analysis software that analyses mass spectrometry data by comparing the spectral counts of a protein between an experimental condition to a negative control. From the SAINT output, we obtained false discovery rates for each protein which indicated the probability of true interaction with TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP. We set FDR thresholds to delineate protein interactors (hits) as such: for a protein to be a unique interactor of one condition, it had to possess an FDR < 0.15 in that condition and an FDR > 0.25 in the opposing condition. Proteins with FDRs < 0.15 in both conditions are deemed shared interactors.

Using these thresholds, we identified 46 significantly enriched proteins, 34 interactors unique to TDP-43^{WT}-GFP and 12 shared between TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP (figure 4B, table 2). These findings indicate that there is an overall loss-of-interactors associated with TDP-43^{Q331K}-GFP.

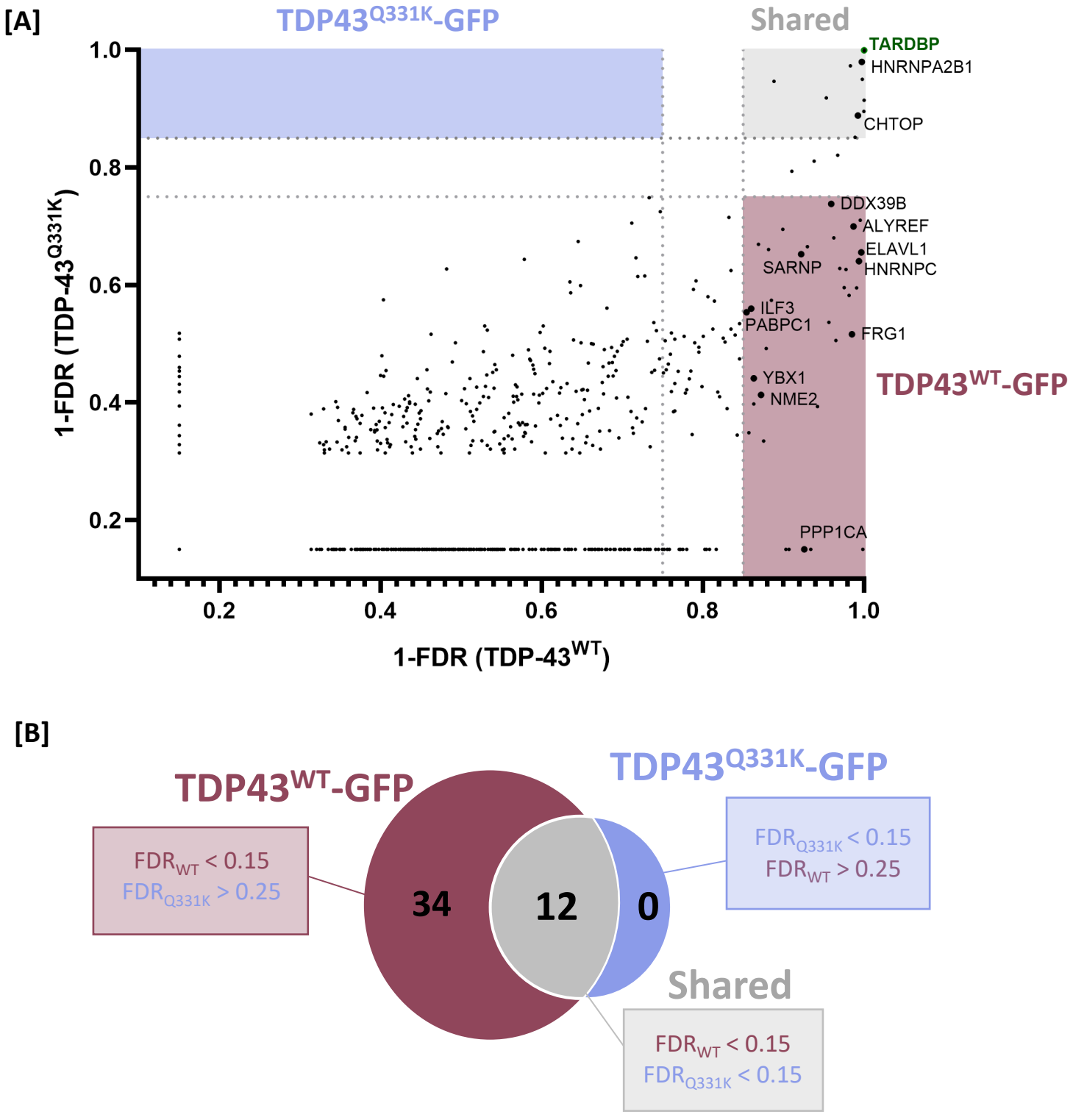


Figure 4: Identification of hits from MS dataset. GFP-IP-MS experiments identified proteins interactors of TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP. **[A] FDR plot of mass spectrometry data.** Each protein is represented by a datapoint and plotted against its FDR of being a true interactor for each experimental condition, represented by shaded areas. TDP-43^{WT}-GFP is represented in maroon, TDP-43^{Q331K}-GFP in light purple and shared hits are shown in grey. **[B] Venn diagram of significant interactors as determined by dynamic thresholds.** Thresholds of FDR < 0.15 (condition 1) and FDR > 0.25 (condition 2) were set to delineate a protein as a unique interactor of condition 1. Shared interactors met a threshold of FDR < 0.15 in both experimental conditions. 34 unique interactors of TDP-43^{WT} (indicated in maroon) and 0 unique interactors of TDP-43^{Q331LK} (indicated in light purple) were identified, suggesting an overall loss of interactions with the mutation. There were 12 shared interactors (indicated in grey) between the two groups.

By performing a functional analysis, we found that these enriched proteins are involved in pathways highlighting TDP-43's many known functions including RNA regulation and transport, splicing, stress granule response and regulation of expression. As expected, many of the hits identified including *HNRNPA0, HNRNPC, HNRNPDL, HNRNPM, HNRNPA2B1* and *SNRNP200* are predominantly found in the nucleus where they play roles in splicing[128].

Pathway analysis of the hits was completed to see whether any of these proteins or functional groups were involved in shared pathways using the Ingenuity Pathway Analysis (IPA) program. IPA allows user to input gene or protein datasets and performs various bioinformatics analyses on the data—we used IPA's pathway analysis function to identify enriched pathways within our data (figure 5). By doing so, we found that some of the hits (*ALYREF, CHTOP, DDX39b, SARNP*) included components of the RNA-protein complex involved in the THO/TREX pathway of nuclear-to-cytoplasmic mRNA export[129-130].

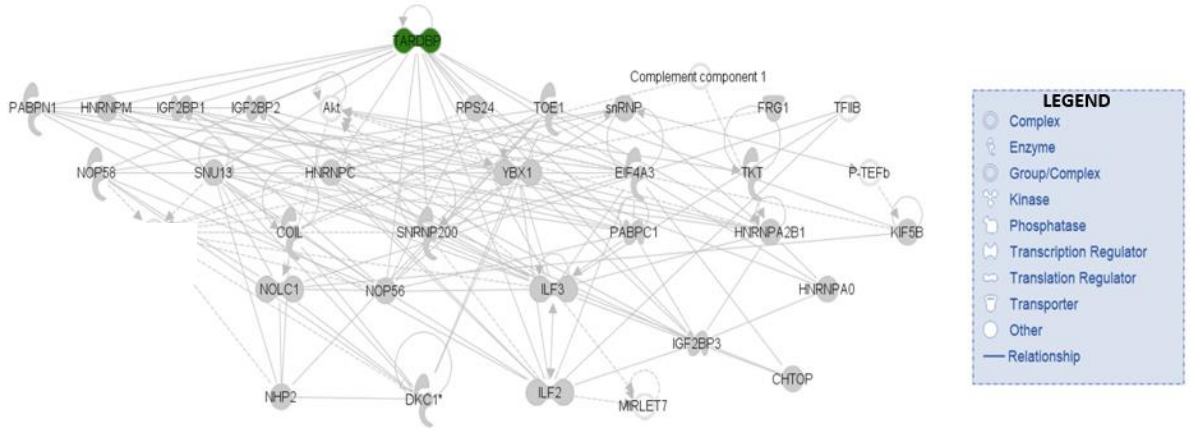


Figure 5: Gene expression, protein synthesis and RNA post-transcriptional modification protein network. One of 3 enriched protein networks identified and mapped using IPA software program. Lines indicate known relationship between proteins and protein type is shown in legend. TDP-43 is shown in green.

In addition to IPA, we browsed the literature for each gene to see whether it has been previously identified in other TDP-43 proteomic/genetic studies or ALS- relevant screens. The hnRNP variants have been previously identified in *wildtype* TDP-43 proteomics studies[109]. Additionally, hits such as *ALYREF, CHTOP, ELAVL1* and *HNRNPC* have been previously identified as proteins of interest in screens completed in yeast cells or drosophila to identify modifiers of TDP-43-mediated toxicity[131-134].

Furthermore, we browsed ALS GWAS and genetic literature; of our protein hits, mutations in the genes *HNRNPA0* and *HNRNPA2B1* have been linked to ALS[135]. Interestingly, some of our other hits have been linked to other neuromuscular diseases that result in pathology or degeneration similar to ALS. *DDX39b* has been implicated in multiple sclerosis, which is an autoimmune disorder of the CNS, resulting in axonal demyelination and neuronal death[136]. *FRG1*, a hit suggested to be involved in epigenetic regulation of muscle differentiation, has been linked to muscular dystrophy and facioscapulohumeral muscular dystrophy 1[137]. *PABPC1*, a protein that co-localizes with stress granules, though has not been shown to have a direct, genetic role in any disease (to our knowledge), has been linked to spinal muscular atrophy and spinocerebellar ataxia as it has been found to be involved in protein complexes with *SMN* and *ATXN2*, respectively[138,139]. *ATXN2* mutations have also been linked to ALS[5].

Finally, though the overlap between transcriptomic and proteomic datasets is limited, we also examined existing ALS transcriptomic datasets, but did not find any shared hits.

Using a combination of these bioinformatic analyses and literature searches, we shortlisted 14 hits for further validation: *ALYREF, DDX39b, ELAVL1, FRG1, HNRNPC ILF3, NME2, PABPC1, PPP1Ca, SARNP, SNRNP200, YBX1, CHTOP* and *HNRNPA2B1* (figure 6). Of these, the latter two are shared interactors between TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP.

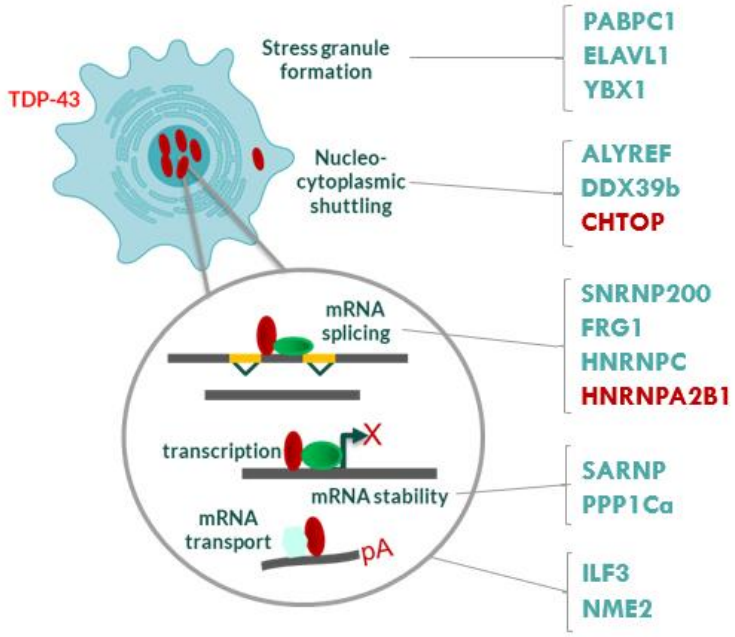


Figure 6: Shortlisted candidates for validation. Shared interactors of WT and mutant TDP-43 have been indicated in red, and interactors unique to WT-TDP-43 have been shown in blue. Interactors have been group based on some functional roles.

7.4 Four protein hits show robust interaction with TDP-43^{WT} with varying decreases in the degree of interaction with mutant TDP-43^{Q331K}

Using the Broad Institute's human ORFeome V8.1 library, we selected and purified DNA from glycerol stocks expressing each of the 12 hits as entry vectors for gateway cloning (table 5). Two of these hits, *NME2* and *SNRNP200* were not present in the library and were purchased in plasmid (from Addgene, for use in cloning) or antibody (table 10, for use in western blots/microscopy) forms. We cloned these hits into destination vectors expressing 3x-Flag tags on the N- or C-terminus of the gene (figure 7). The location of the 3x-Flag tag was determined by examining existing studies to select the terminus for each protein that would have minimal or no impact on protein function or interaction (table 4).

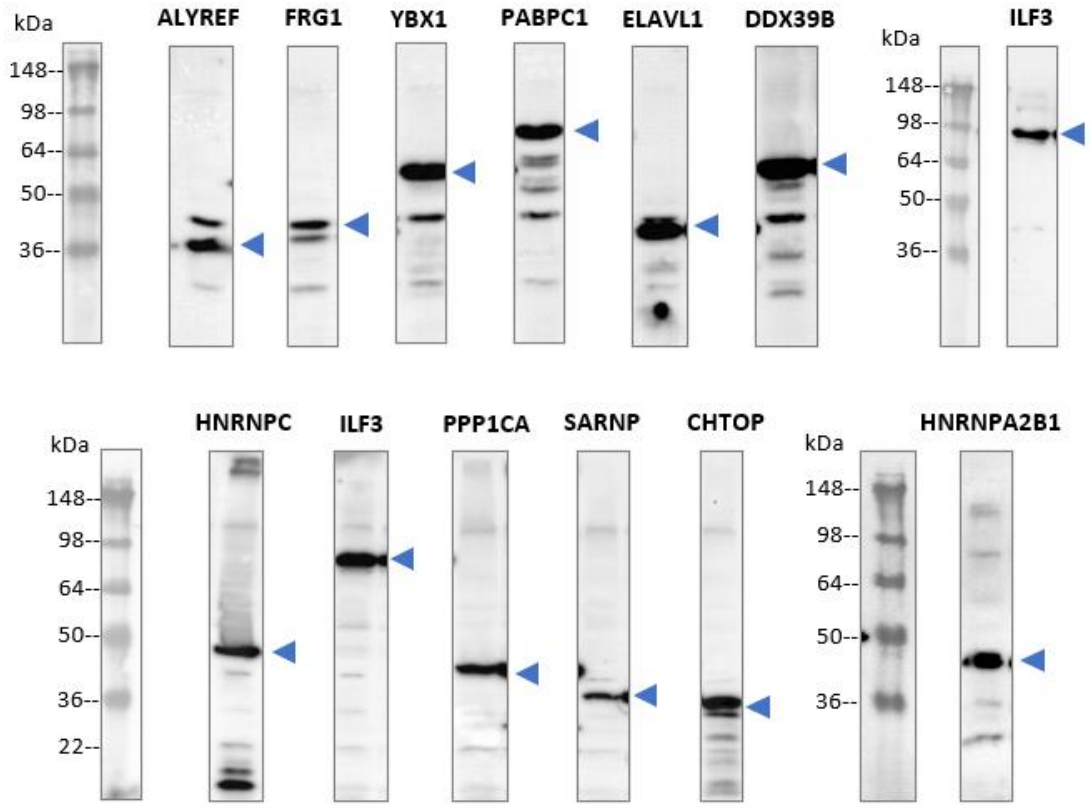
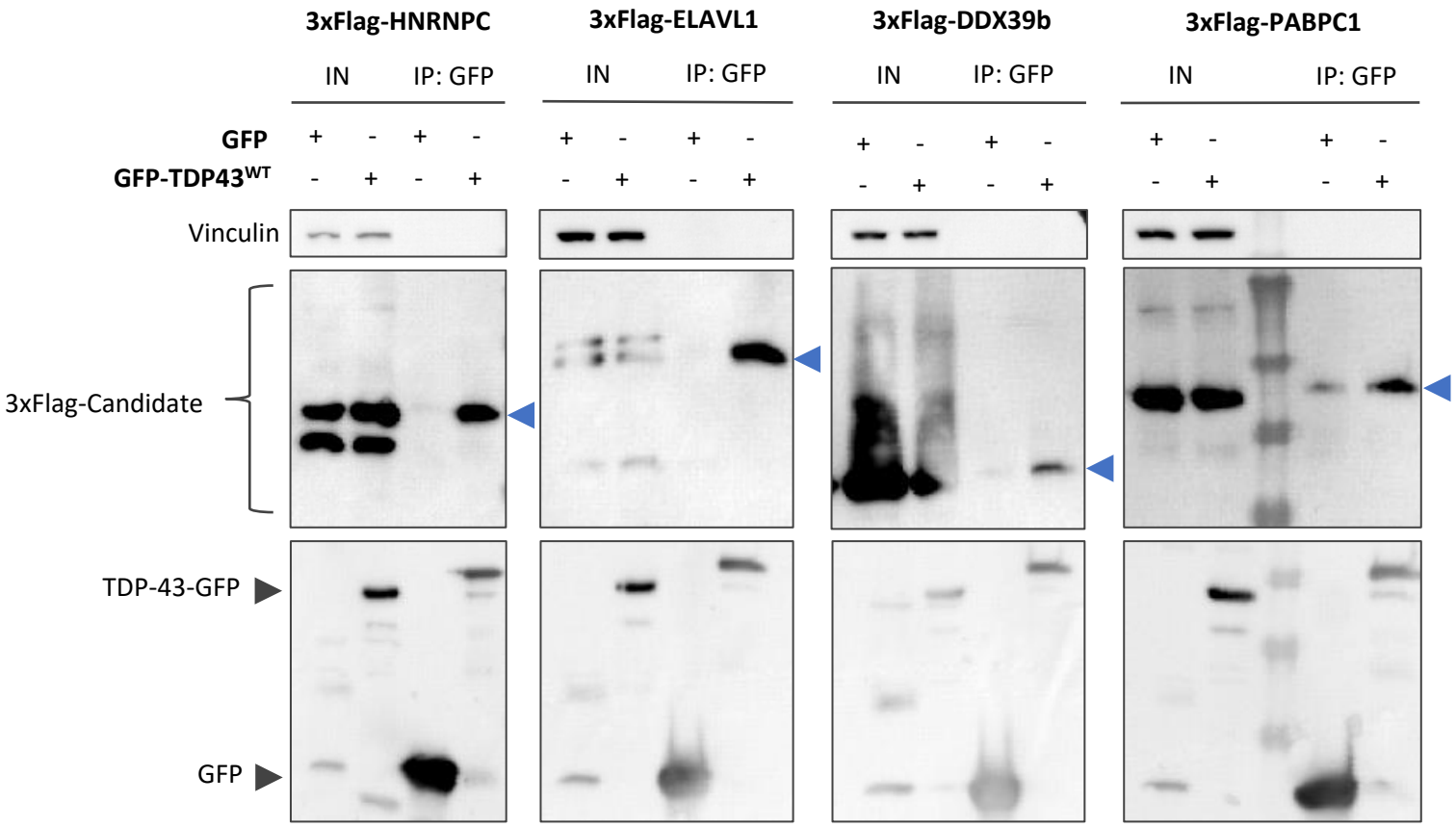


Figure 7: Test expression of 3x-Flag-hit plasmids. 3x-Flag-Hit plasmids were generated using various cloning approaches and transfected with ePEI reagent into HEK293T cells. Cells were incubated at 37ºC for 48 hrs to allow for expression and then lysed in 1X NETN lysis buffer. Lysate was collected and processed for immunoblot. Protein ladders have been grouped with lanes from its respective gel and molecular weight (kDa) is indicated. Immunoblots have been probed using rat α-Flag (1:1,000). Blue arrows indicate expression of 3x-flag-tagged protein of interest. Presence of multiple bands may represent higher molecular weight oligomers or cleavage products.

To validate interaction with TDP-43^{WT}-GFP, we co-transfected these constructs into HEK293T cells with plasmids expressing GFP-tagged TDP-43^{WT} or unconjugated GFP as a negative control. Similar to the initial screen, we performed GFP-immunoprecipitation to pull down TDP-43^{WT}-GFP interactors and then subsequently performed an immunoblot to determine whether these 3x-flag-tagged-hits were present in the pulldown (figure 8A, S1). Of the 14 hits, 4 proteins, encoded by the *ELAVL1, DDX39b, PABPC1, HNRNPC* genes, showed strong, robust interaction with TDP-43^{WT} and showed a consistent and high frequency of validation across several GFP-immunoprecipitation-WB experiments (figure 8B).



Relative Percentage of Validation via IP-WB	
Gene of Interest	N (tested)
ELAVL1	4
DDX39b	4
PABPC1	4
HNRNPC	4
ILF3	3
ALYREF	4
CHTOP	4
PPP1Ca	4
SARNP	4
YBX1	4
FRG1	4
HNRNPA2B1	3
NME2	4

Figure 8: Validation of hit interaction with TDP-43^{WT}-GFP.

[Top] Representative WBs showing co-overexpression GFP-IP experiments where GFP-TDP-43^{WT} constructs were co-transfected with 3xFlag-hits (ELAVL1, DDX39b, PABPC1, HNRNPC) in HEK293T cells, incubated to express for 48 hours. Blue arrows indicate expression of 3xFlag-hit. Top WB: probed using vinculin antibody. Middle WB: probed using Flag antibody. Bottom WB: Probed using GFP antibody.

[Left] Relative percentage of validation of 3x-flag-hits using GFP-IP and subsequent WB. Right column indicates total number of times hit of interest was tested.

Next, we tested whether the presence of the Q331K mutation disrupts these interactions using the same validation paradigm. We did not observe significant differences in the degree of interaction between the hit candidates *HNRNPC, ELAVL1, PABPC1* and GFP-expressing TDP-43^{WT} or TDP-43^{Q3331K} (figure 9). 3xFlag-*DDX39b* did not show expression at the correct molecular weight on the immunoblot and was excluded from data analysis.

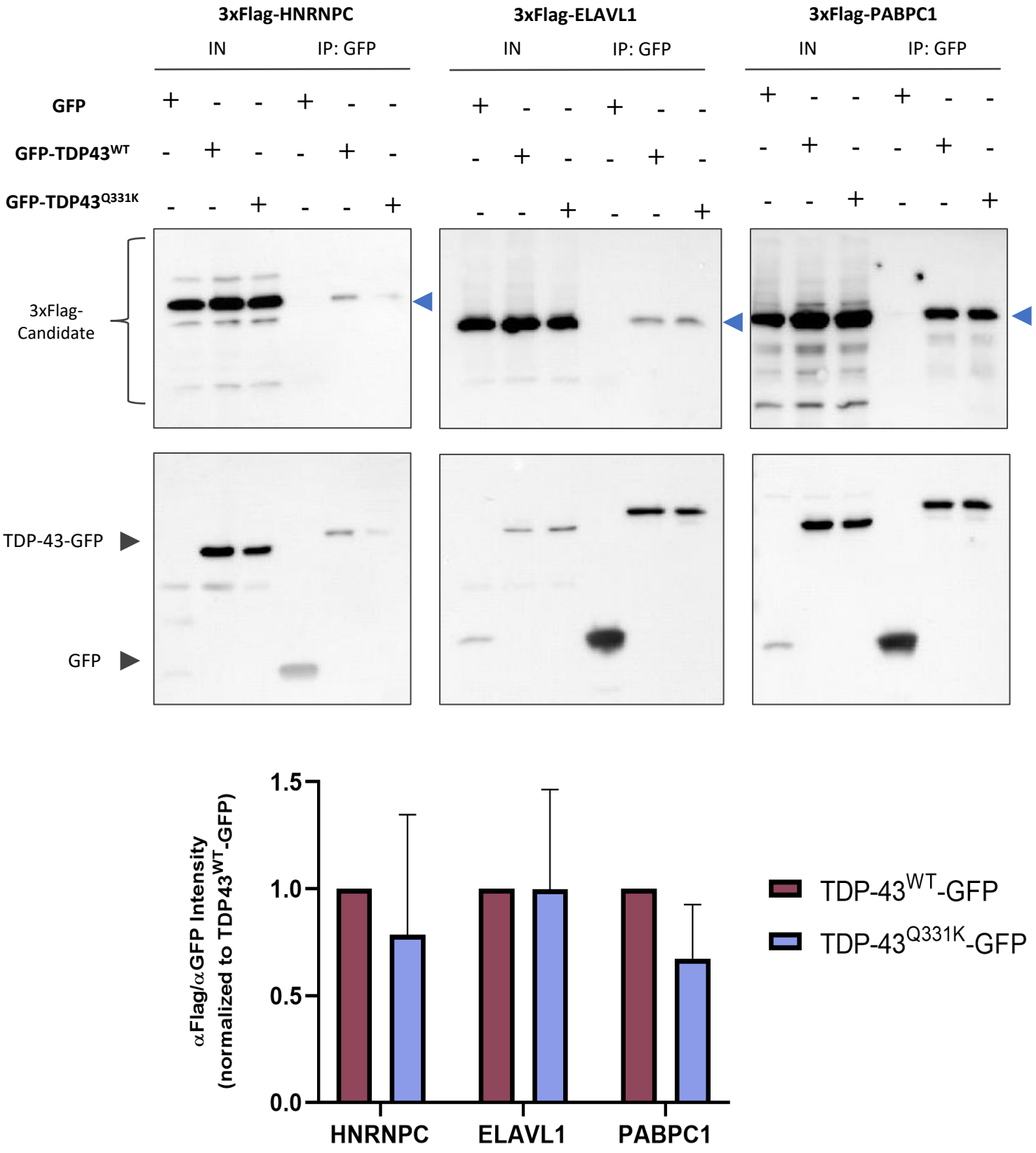


Figure 9: Assessing loss of interaction in TDP-43^{Q331K} of top hits using co-overexpression approach. [Top] WB showing GFP-IP of co-overexpression of 3xFlag-tagged HNRNPC, ELAVL1, PABPC1 and GFP (negative control), GFP-TDP-43^{WT} or GFP-TDP-43^{Q331K} in HEK293T WT cells, incubated over 48 hr expression interval. Blue arrow indicates pulldown of 3xFlag-candidate. Top WB: probed using Flag antibody. Bottom WB: Probed using GFP antibody. **[graph]** Quantification of 3xFlag-hit pulldown using GFP-IP. Quantification completed by standardizing relative flag signal to relative GFP signal, normalized to GFP-control. Bars represent mean of experiments. Error bars indicate standard deviation. (n = 3).

Given that TDP-43's expression is tightly regulated, we hypothesized that the co-overexpression nature of our experiment may have 'forced' interactions to occur that may not reflect the true biological nature of cells. We attempted to circumvent this by using a similar experimental paradigm that substitutes the GFP-TDP-43 expressing constructs for the knock-in cells. We overexpressed each hit in the endogenously expressing TDP-43^{WT}- and TDP-43^{Q331K}- GFP knock-in cells and then performed a GFP-immunoprecipitation-western blot (figure 10A). Contrary to our dual overexpression system, but similar to our IP-MS data, we observed a significant decease in the degree of hit interaction with TDP-43^{Q331K}-GFP for both DDX39b and HNRNPC conditions (figure 10B). Additionally, we also observed a decrease in the degree of interaction with the TDP-43^{Q331K}-GFP KI cells when ELAVL1 was singly overexpressed in two of the three replicates. Statistical analysis of this data indicated that the third replicate is likely an outlier, suggesting ELAVL1's interaction with TDP-43 is reduced or lost in the case of the Q331K mutation. In two of three biological replicates, α-flag-probed immunoblots of the PABPC1 GFP-IPs indicated successful transfection of the 3x-Flag construct through its expression in the input sample, however, did not show the presence of any flag-tagged protein in the pulldown lanes, rendering the results from this set of IPs to be inconclusive. Together, these findings suggest that there are varying decreases in the degree of hit interaction with TDP-43^{Q331K}, where DDX39b and HNRNPC show significant decreases, and upon removal of an outlier, ELAVL1 shows a trending decrease. The results for PABPC1 were inconclusive and may require some further optimization of the GFP-IP conditions.

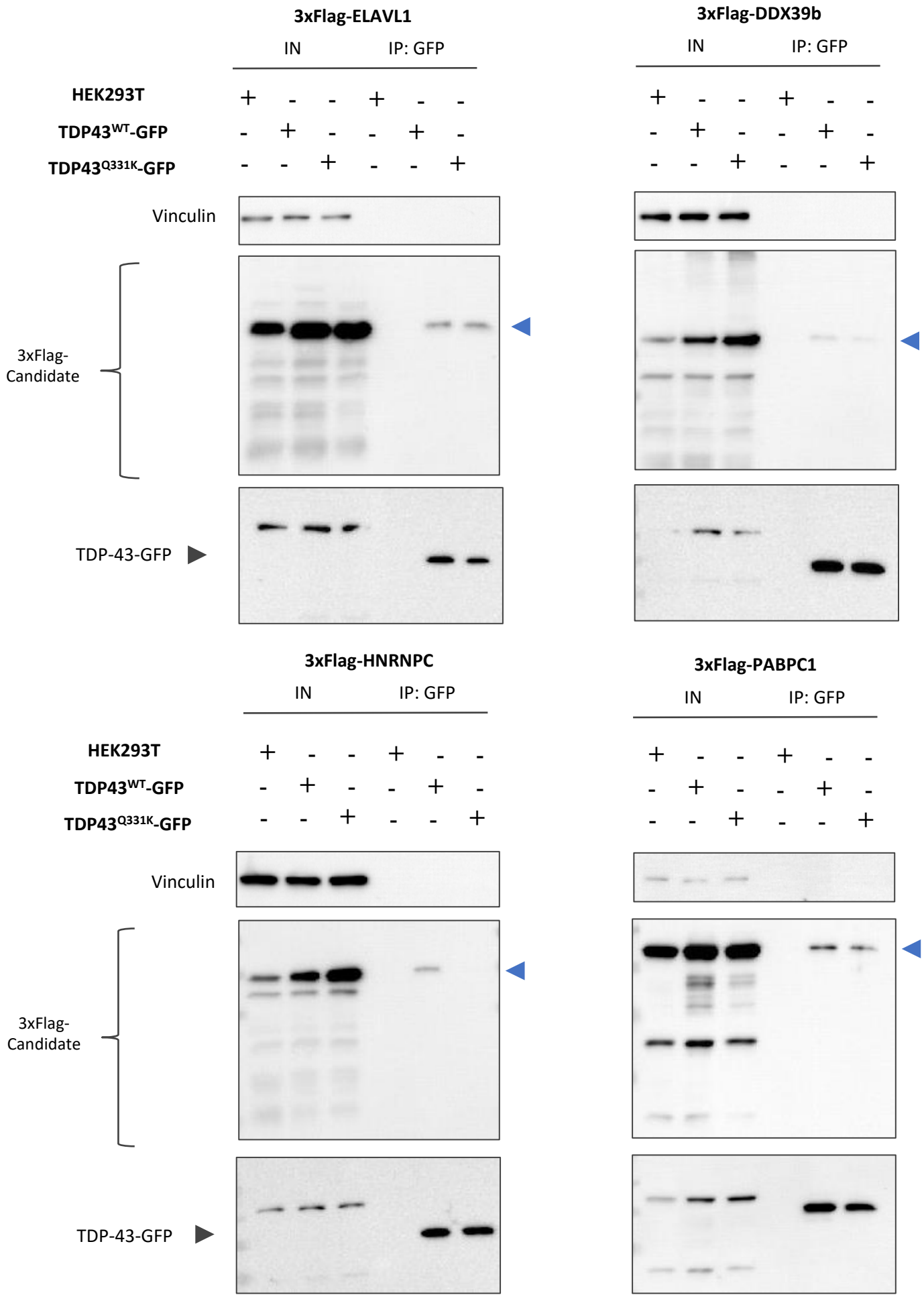


Figure 10 (caption on next page)

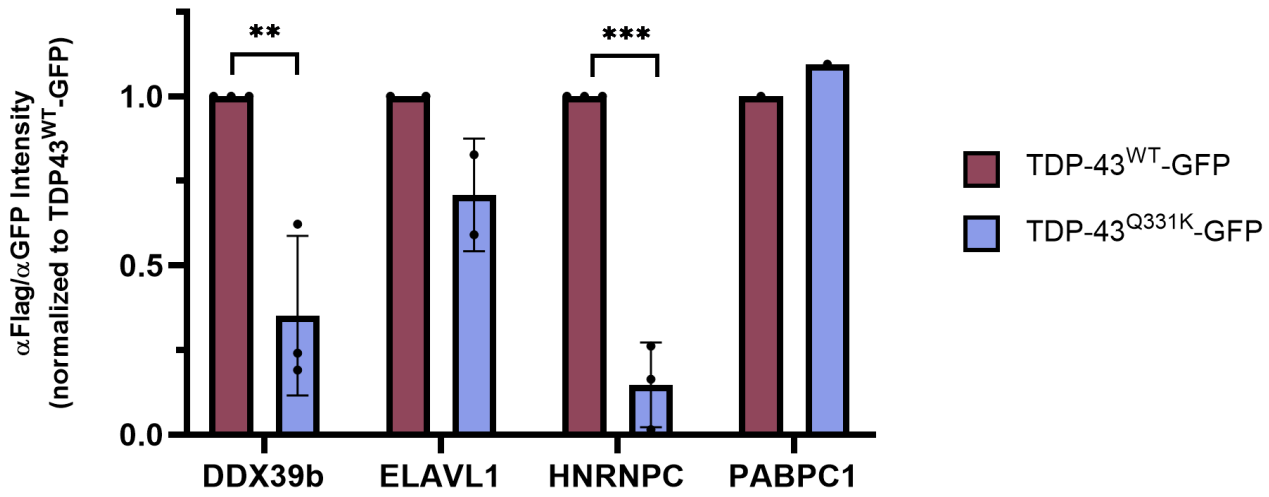


Figure 10: TDP-43^{Q331K} results in varying decreases in the degree of interaction with 3xFlag-Hits. [Previous] WB showing GFP-IP of overexpression of 3xFlag-tagged hits of interest. 3xFlag-tagged- HNRNPC, ELAVL1, PABPC1 in GFP-TDP-43^{WT} or GFP-TDP-43^{Q331K} knock-in cells, incubated to express for 48 hours. Blue arrow indicates pulldown of 3xFlag-candidate. Top WB: probed using vinculin antibody. Middle WB: probed using flag antibody. Bottom WB: Probed using GFP antibody. **[graph] Quantification of 3xFlag-hit pulldown using GFP-IP.** Quantification completed by standardizing relative flag signal to relative TDP-43-GFP signal, normalized to GFP-TDP-43^{WT} TDP-43 signal intensity. Bars represent mean of experiments. Error bars indicate standard deviation. (n = 3 for DDX39b and HNRNPC. n = 2 for ELAVL1 due to removal of outlier, and n = 1 for PABPC1 due to absence of αFlag signal in IP lanes on WB). Groups were analyzed using multiple unpaired t-tests with post-hoc analysis. **$p < 0.01$, ***$p < 0.001$.

7.5 Knockdown of top hits results in increased nuclear localization of TDP-43

To assess the effects of hit knockdown on TDP-43 localization, we generated constructs expressing two shRNAs for each of the 14 top hits (figure S3). The knockdown efficiency of these shRNAs was tested at least twice via qPCR, and shRNAs against the top four validated hits underwent additional qPCR analysis to confirm their effectiveness (figure 11).

This qPCR data shows a significant reduction in gene expression compared to a non-targeting luciferase control for three shRNAs (against *DDX39b* and *PABPC1*), and trending, but non-significant, decreases in gene expression for all other shRNAs against the top four hits.

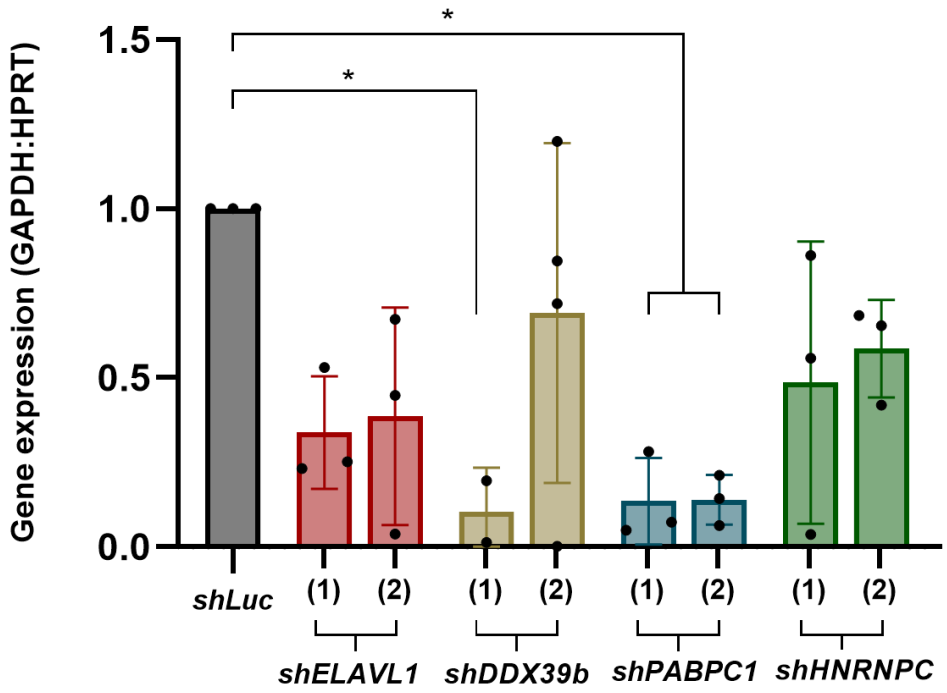


Figure 11: Knockdown efficiency of shRNAs against top hits as assessed by qPCR. (n = 3 unless otherwise indicated by datapoints) Experiments have been plotted as individual data points. Bars represent mean of experiments with standard deviation error bars. Gene expression (2^{-ddCT}) of shRNA target is normalized to average of housekeeping genes (GAPDH and HPRT). Outlier in shDDX39b (2) has been removed due to discrepancies in qPCR reads. One-way ANOVA with post-hoc test was performed to assess for significant knockdown. *indicates $p < 0.05$.

Next, to assess whether KD of our top hits alters TDP-43's subcellular localization, we used an experimental paradigm in which we transfected HEK293T cells with two shRNAs against each hit and let the cells

express the constructs over a 72-hour interval to allow for target gene knockdown. We then imaged the cells using fluorescent microscopy (figure 12, S4) and used an image analysis software called Cell Profiler to assess the degree of TDP-43 nuclear-to-cytoplasmic mislocalization. The Cell Profiler pipeline used to analyze the images was programmed such that it identified nuclei using information from the DAPI channel and cell bodies using YFP channel data, as primary and secondary objects respectively (figure S4). The cytoplasm was determined by subtracting nuclei from its parent cell body. αTDP-43 signal intensity was measured, and the cytoplasm-to-nuclear intensity ratio was calculated for each YFP-positive cell per image. We completed this experimental paradigm in three individual replicates using cells from different passages. However, one of these biological replicates resulted in low transfection efficiency and was excluded from the final data analysis. Additionally, we originally aimed to assess the effects of both hit OE and KD on TDP-43 localization within cells. However, we were unable to optimize staining of the 3x-Flag tag to achieve a good signal-to-background ratio which would be required for quantification of the microscopy images, and this aspect of the study was excluded.

At baseline levels, which were represented by the non-protein targeting shLuciferase treated cells, we observed a larger proportion of TDP-43 localized in its normal nuclear locale, as expected. Upon shRNA-mediated knockdown of our four top hits, we observe a significant decrease in the cytosol-to-nuclear ratio of TDP-43 compared to the shLuciferase control. Together, these findings suggest that there could be an increase or shift in the nuclear localization of TDP-43 with decreased expression of these four hits.

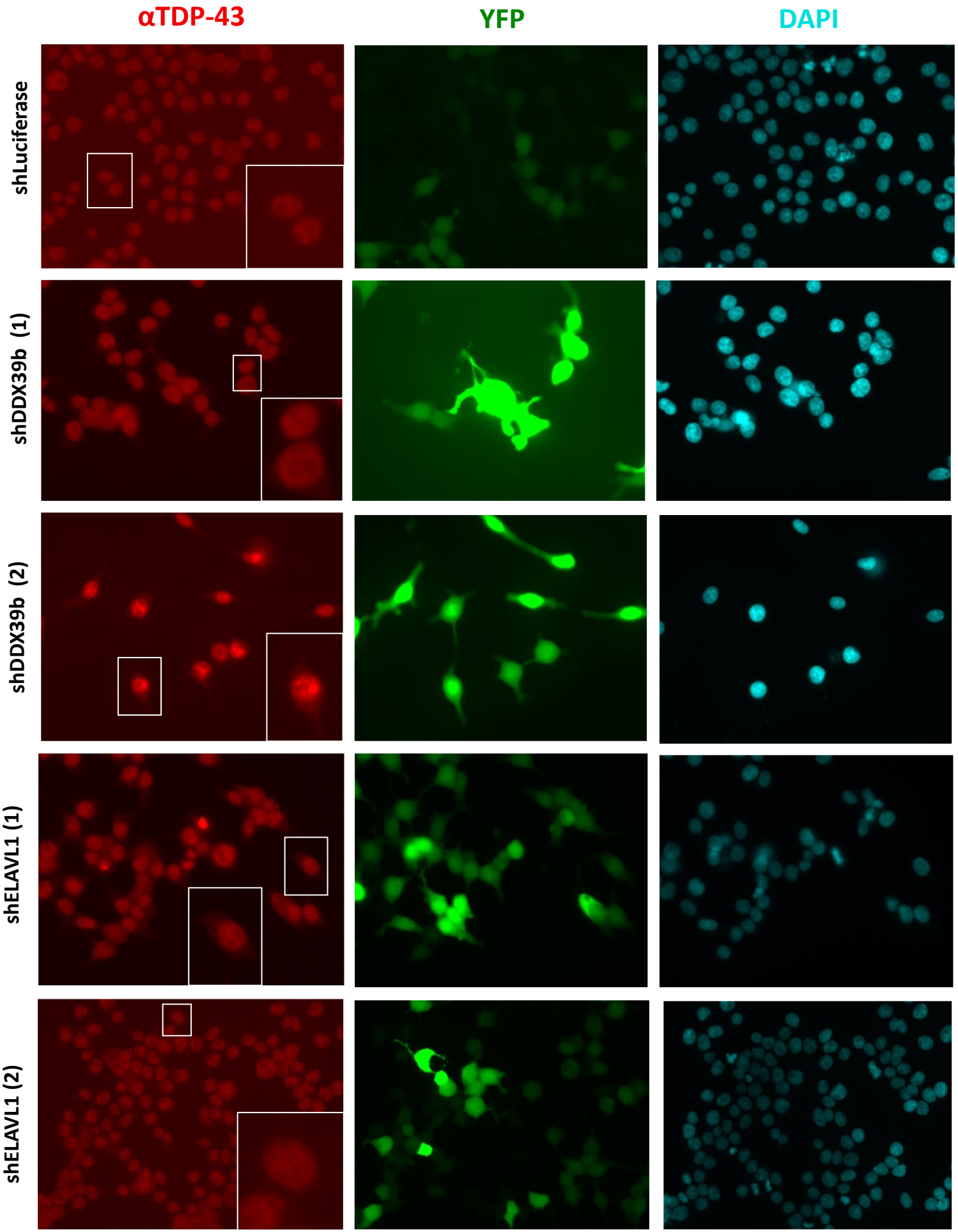


Figure 12 (caption on page 50)

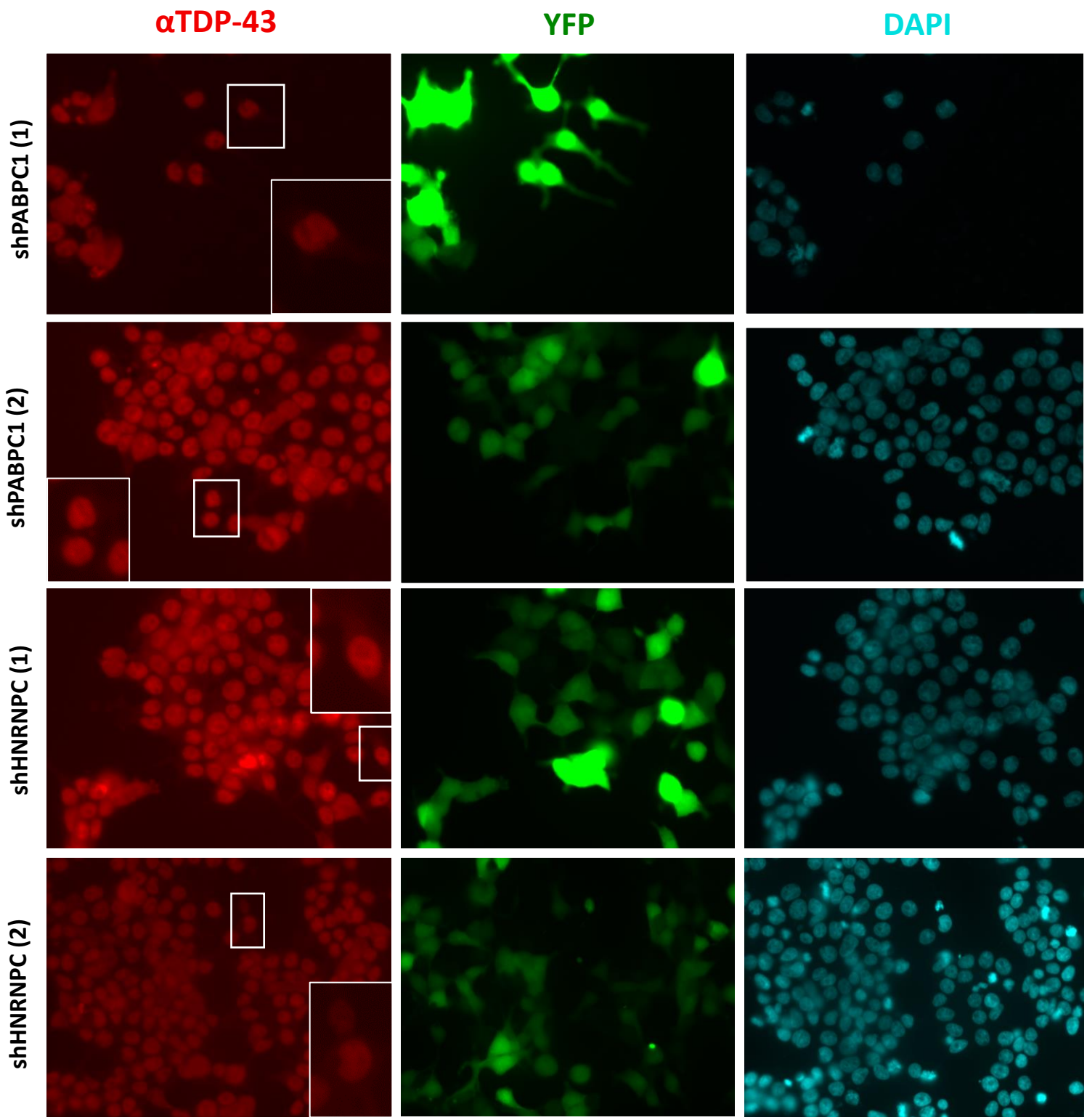


Figure 12 (caption on page 50)

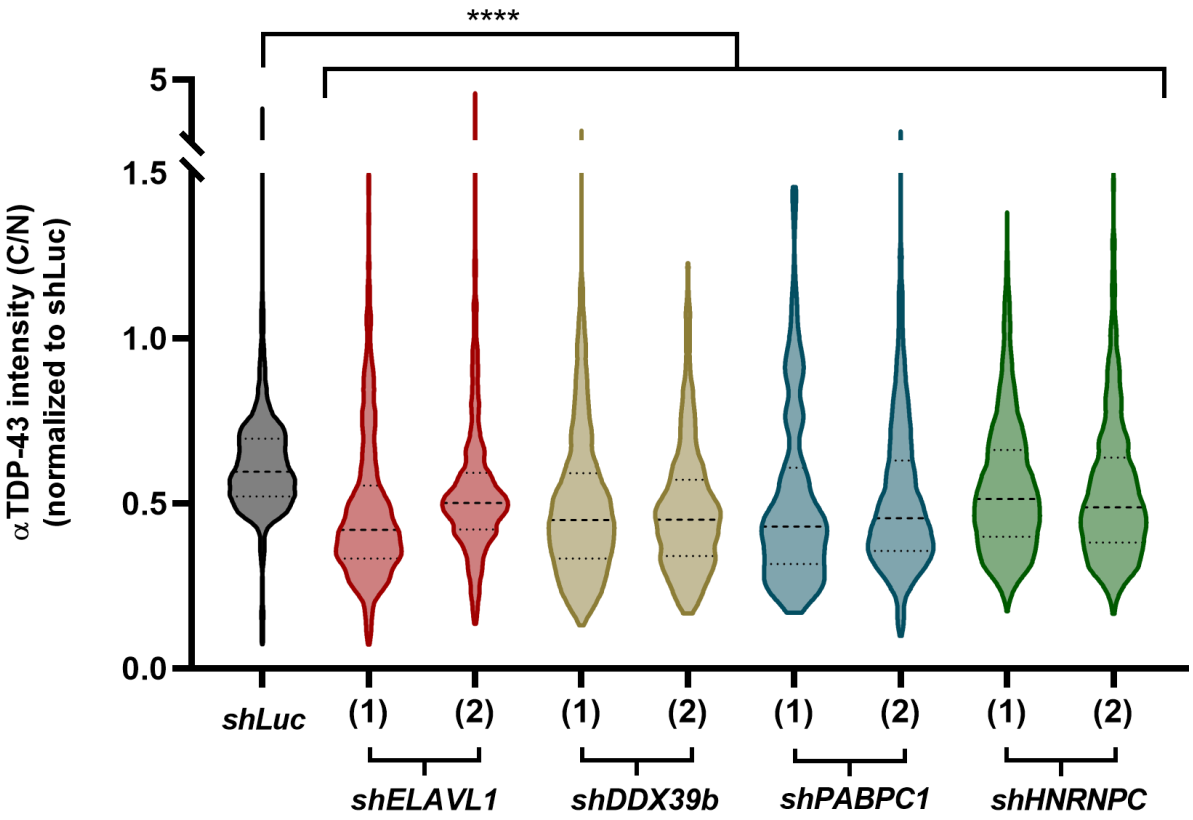


Figure 12: Assessing TDP-43 localization upon shRNA-hit knockdown via microscopy. Representative immunofluorescent microscopy images. HEK293T cells were transfected with shLuciferase, shDDX39b, shELAVL1, shHNRNPC and shPABPC1, with expression for 72 hours. DAPI (blue), TDP43 (red) and EGFP (green) channels of images is shown. Images have been cropped and enlarged to highlight TDP-43 subcellular localization. Red channel images brightness and contrast has been adjusted by Original images have been included in supplemental figures. **[graph] Violin plot showing distribution of C/N ratios measured per cell per image by Cell Profiler.** Grey line indicates mean αTDP-43 intensity (C/N) for luciferase. 100-200 cells were counted per sample per replicate. Graph shows individual cell counts from two combined microscopy experiments, third experiment was completed but removed due to low transfection number and cell counts, resulting in discrepancies in C/N ratio. One-way ANOVA test was completed with post-hoc analysis. **** indicates $p < 0.0001$.

7.6 Knockdown or overexpression of top hits does not affect TDP-43's splicing or RNA binding abilities

Finally, to assess whether a gain or loss of TDP-43:hit interaction has a functional consequence on splicing which is a well-defined function of RNA binding proteins like TDP-43, we used a RT-PCR based splicing assay. TDP-43 loss of function studies done by *De Conti et al*. in HEK293 and HeLa cells to identify transcripts that are differentially spliced upon TDP-43 knockdown have identified a number of genes that show altered splicing, including *POLDIP3* (polymerase delta interacting protein 3)[140]. Upon siRNA-mediated knockdown of TDP-43, there is increased exon 3 exclusion of *POLDIP3*, resulting in an upregulation of its variant 2 transcript (figure 13A). Similarly, a paper by *Shiga et al.* found that the RNA binding ability of TDP-43 is necessary for exon 3 inclusion of *POLDIP3*, and that motor cortex, spinal cord and spinal motor neuron samples from ALS patients show an upregulation of *POLDIP3* variant 2 mRNA, providing support of the TDP-43 loss-of-function hypothesis[141].

The 3x-flag-hit overexpression constructs and shRNAs that were generated were transfected into HEK293T cells, and then incubated for 72 hours to allow for expression. After this interval, the cells were pelleted to isolate RNA and synthesize cDNA, which was used as the input for PCR amplification of *POLDIP3* using primers (see table 9) described in the *De Conti et al.* study. PCR samples were run on DNA PAGE gels, and exon 3 exclusion was calculated by quantifying the resultant bands. We observed a significant increase in exon 3 exclusion upon shRNA-mediated knockdown of TDP-43, resulting in a loss of its function (figure 13B). However, shRNA-mediated knockdown or overexpression of the top hits did not have a significant affect on exon 3 exclusion (figure 13B and 13C), suggesting that these hits may not be involved in mediating TDP-43's splicing or RNA binding abilities.

Interestingly, we observe a trending increase in exon 3 exclusion in both the TDP-43^{WT} and TDP-43^{Q331K} GFP KI cell lines compared to HEK293T cells (figure S6), suggesting that the addition of the GFP tag may result in some loss of TDP-43 function. However, this difference is not significant between cell types.

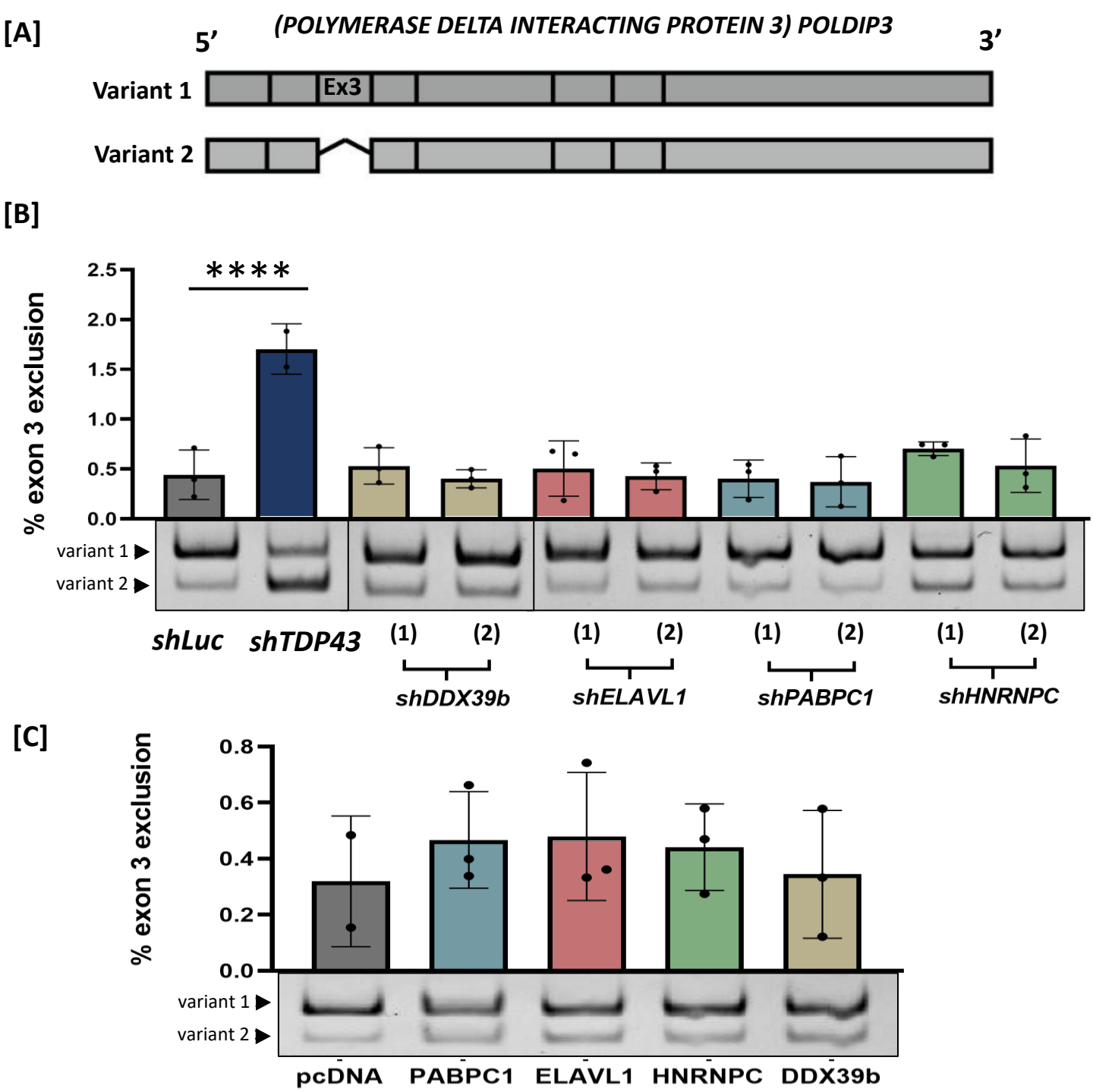


Figure 13: RT-PCR based splicing assay shows no significant TDP-43 loss of function with knockdown or overexpression of top hits. [A] *POLDIP3* mRNA transcript 1 and 2. Variant-2 mRNA is upregulated with TDP-43 loss of function results due to exon 3 exclusion from the full length transcript. **[B] shRNA-mediated KD of TDP-43 and top four hits.** Graph showing % of *POLDIP3* exon 3 exclusion, quantified from TAE PAGE gel bands (shown below graph) for shLuc, shTDP43, shDDX39b, shELAVL1, shPABPC1 and shHNRNPC. shRNA mediated knockdown of TDp-43 results in a significant upregulation of variant-2 mRNA. Knockdown of top hits does not significantly affect TDP-43 function. **[C] OE of top hits.** Graph showing % of *POLDIP3* exon 3 exclusion, quantified from TAE PAGE gel bands (shown below graph) for 3x-flag-tagged: PABPC1, ELAVL1, HNRNPC and DDX39b and pcDNA. Overexpression of top hits does not significantly alter TDP-43 function.

Data was analyzed using one-way ANOVA with post-hoc analysis for multiple comparisons. Bars represent mean of experiments. Error bars indicate standard deviation. All experiments are an n of 3, unless otherwise indicated by datapoints. **** $p < 0.0001$.

8. DISCUSSION

Summary of findings

Here, we show the generation and characterization of isogenic, stably expressing knock-in cell lines in which wild-type and ALS-causing mutant Q331K, TDP-43 are expressed at endogenous levels. In this model, we observed that the ALS-causing TDP-43^{Q331K}-GFP mutation results in reduced expression of the microtubule regulator *STMN2*, which suggests TDP-43 loss-of-function. In addition to loss of function, we have also observed significant increases in TDP-43^{Q331K}-GFP cytoplasmic localization. Using these knock-in cells as an experimental tool, we identified TDP-43 protein-protein interactions using mass spectrometry and found that the TDP-43^{Q331K}-GFP mutation leads to an altered interactome, with an overall loss of TDP-43 interactors. We shortlisted 14 hits whose native interactions with TDP-43 are lost upon introduction of the Q331K mutation and validated them using IP-WB. of these, 4 proteins (*DDX39b, ELAVL1, HNRNPC* and *PABPC1*) showed a high percentage of validation with TDP-43^{WT}-GFP using our IP-WB approach, and some of these were significantly decreased with TDP-43^{Q331K}-GFP. shRNA-mediated KD of these four proteins resulted in increased TDP-43 nuclear localization, suggesting that there is a functional effect associated with loss-of TDP-43-hit interaction. However, we do not observe significant TDP-43 loss of function upon hit overexpression or knockdown using an RT-PCR based splicing assay. This work suggests that *TARDBP* ALS-linked mutations alter the TDP-43 interactome, generally resulting in a loss of interactors, which may have functional consequences on its subcellular localization.

Mass spectrometry data supports the use of TDP-43-GFP KI cell model for interactome mapping and shows that TDP-43^{Q331K} mutation results in altered interactome with loss-of-interactors.

We observed an overall loss in many TDP-43-protein interactions that are involved in pathways responsible for splicing, RNA post translational modifications, gene expression and protein synthesis, suggesting that these pathways or protein-protein interactions are altered with TDP-43 ALS-linked mutations. Of the 46 proteins that met our significance thresholds, several proteins including hnRNP family members, SNRNP200, EIF4A3, ILF3, PTBP1, IGF2BP1/2/3, among others have been identified in other TDP-43 proteomics studies[109-111]. Additionally, members of proteins involved in nucleo-cytoplasmic shuttling via

the THO/TREX pathway and in the stress granule response were also enriched within our dataset—both of which are cellular functions that have been shown to be impaired or dysfunctional in ALS[129-131,142].

Our mass spectrometry data indicates that the abundance TDP-43^{WT}-GFP and TDP-43^{Q331K}-GFP present between the samples was relatively equivalent, and as such, any decrease or loss we observe in interactors is likely due to structural or functional differences between the two proteins variants, and not due to reduced protein levels in the input sample. Of these 46 hits that met our significance thresholds, 14 were selected for validation via GFP-immunoprecipitation-western blot. From these validation experiments, DDX39b, ELAVL1, HNRNPC and PABPC1, are RNA-binding proteins that showed consistent and robust interaction with TDP-43^{WT}-GFP and varying degrees of decreases in interaction with the TDP-43^{Q331K}-GFP mutation.

HNRNPC

HNRNPC is a ubiquitously expressed RNA binding protein that, like TDP-43, is a member of the heteronuclear ribonucleoprotein family of proteins. It is primarily nuclear-localized and suggested to play a role in pre-mRNA processing[143]. While other members of this protein family, including HNRNPA0 and HNRNPA2B1 have been linked to ALS via genetic mutations, HNRNPC has not been directly linked to ALS pathogenesis to our knowledge. However, this protein has been previously identified in a TDP-43 interactome study and was validated using an IP-WB experimental paradigm, providing support for our findings and validation strategy[110].

ELAVL1

ELAVL1 is an RNA binding protein that binds and stabilizes mRNA and regulates splicing. Many of its known protein-protein interactions are RNA-dependant, including some of which have been identified in our screen including: IGF2BP1, IGF2BP2, IGF2BP3, ILF3, PABPC1, PTBP2 and YBX1[144]. ELAVL1 was also identified in a functional yeast screen by *Couthouis et al*. which aimed to identify RNA binding proteins that have similar properties as the ALS-linked proteins, TDP-43 and FUS[145]. Its expression in yeast resulted in the formation of highly toxic cytoplasmic aggregates and other studies have shown that ELAVL1 has the ability to bind and regulate TDP-43's expression and toxicity[144,146].

PABPC1

PABPC1 is an RNA binding protein that is found in stress granules and at the sites of high translational activity[147]. It is localized primarily in the cytoplasm and thought to shuttle into the nucleus through passive diffusion.

A study by Hyung-Jun *et al.* performed a genome-wide yeast screen to identify proteins that associated with stress granules in conditions of TDP-43 toxicity. They found that eIF2α phosphorylation, an event that is indicative of the formation of stress granules, becomes upregulated with TDP-43-associated neurodegeneration[148]. They identified multiple genes that act as modifiers of toxicity, and these genes all linked to PAB, the drosophila ortholog of PABPC1[146,147]. Further characterization of this hit found that the protein is required to drive TDP-43 mediated toxicity in drosophila. Furthermore, PABPC1 was also shown to be mislocalized and accumulated into inclusions in spinal cord motor neurons derived from ALS tissue[147,148].

Stress granules comprise a key aspect of ALS research given that they have been linked to the disease through several mechanisms. Firstly, mutations in the *TIA1*, encoding for the Tia1 cytotoxic granule-associated RNA binding protein cause ALS[149]. Many other ALS-linked proteins including TDP-43, FUS, and dipeptide repeats of C9ORF72 have been shown to regulate, colocalize with or alter the dynamics of stress granules[150,151]. Furthermore, studies have shown that prolonged stress causes TDP-43 proteinopathy, similar to that seen in ALS patients[1]. Although we were not able to validate the loss of PABPC1:TDP-43^{Q331K} interaction as we were not able to confirm expression of the 3xFlag-PABPC1 construct in two of our three GFP-IP experiments, the literature suggests that stress granule pathways or complexes may be altered in ALS[150], and further validation and studies are needed to better understand our findings.

DDX39b

DDX39b is an essential splicing factor and component of the TREX complex in the THO/TREX pathway which is responsible for binding spliced mRNA and mediating its transport out of the nucleus, thereby coupling mRNA processing and nuclear export[152]. DDX39b, along with several other members of the

TREX complex, were identified in an interactome study by *Boehringer et al*. that aimed to identify protein interactors of WT and ALS-linked, matrin3, encoded by the *MATR3* gene[153].

The THO/TREX Pathway of Nuclear Export

Some of our top hits included components of the THO/TREX pathway which includes proteins encoded by the *ALYREF*, *SARNP*, *CHTOP* and *SNRNP200* genes, identified in our study. Other proteins which comprise the protein:mRNA complex involved within this pathway include members of the Thoc family of proteins, the nuclear export protein Nxf1:p15, shuttling SR proteins, and members of the exon junction complex, among others[130,131]. However, these proteins were not identified as significant interactors in our mass spectrometry experiments. Alterations in or impairments in mRNA nuclear export pathways such as the THO/TREX pathway in the TDP-43^{Q331K} knock-in cells could affect genes whose transcripts are shuttled using this transport system, thereby altering the proteome as we see with the TDP-43^{Q331K}-GFP mutant.

Loss of TDP-43-hit interaction results alters TDP-43 subcellular localization

We observed increased TDP-43 nuclear localization upon shRNA mediated KD of our four hits of interest, suggesting that there is a functional consequence associated with loss of TDP-43-hit interaction that may alter its localization. Given that TDP-43 is an RNA binding protein that modulates gene expression by binding to and regulating transcripts of other proteins, a potential mechanism for this observation may be that the transcripts of these top hits are natively regulated by TDP-43, and a downregulation of their protein expression results in compensatory mechanisms that results in increased TDP-43 nuclear localization, where its return to the nucleus can upregulate these transcripts and bring hit protein levels closer to baseline.

Splicing assay shows no significant loss of TDP-43 function upon OE or KD of top hits

Using a splicing assay previously described in the literature and used to assess TDP-43 loss of function, we tested our OE and KD constructs to see whether a gain or loss of our top hits would have a functional consequence on TDP-43. This RT-PCR based splicing assay has been characterized to show a sensitive readout of TDP-43 loss-of-function in HEK293 cells[140], where TDP-43 loss of function is attributed to increased exon 3 exclusion of *POLDIP3*, a gene whose transcript is regulated by TDP-43[140,141].

As a proof of concept, we were able to show that shRNA-mediated KD of TDP-43 results in increased exon 3 exclusion of *POLDIP3*, indicative of TDP-43 loss of function. Furthermore, our experiments using HEK293T cells showed that the percentage of exon 3 exclusion of *POLDIP3* lies within the confidence intervals outlined in the *De Conti et al*, using the same cell strain and primers[140], further supporting the validity of our experimental paradigm and findings. We did not observe any significant differences upon overexpression or shRNA-mediated KD of our top hits, suggesting that our top hits may not be in involved in mediating TDP-43's splicing or RNA binding abilities[141]. Furthermore, we observed a trending increase in exon 3 exclusion, but no significant differences between HEK293T cells and the GFP KI cells, suggesting that there may be TDP-43 loss of function that is mediated by the GFP-KI—and further studies using alternative structural or functional readouts may provide more conclusive insight. Studies have shown that the cellular transcriptome is altered in ALS patients and given the overall loss of interactors identified in our mass spectrometry experiments, it is still likely that we may identify transcriptomic changes in our GFP-KI cells compared to HEK293T cells or with OE or KD of our top hits (simulating a gain or loss of TDP-43:hit interaction) using other assays or transcriptomic readouts.

Possible avenues facilitating pathology in TDP-43^{Q331K}

Given that the hits, DDX39b, ELAVL1, and HNRNPC are primarily found in the nucleus, and we have observed increased TDP-43 cytoplasmic localization in the mutant Q331K cells, we would expect to see a loss or reduction in protein binding to TDP-43^{Q331K}-GFP, which we have observed using IP-WB. In addition, studies have found that mutations in TDP-43's CTD, such as the TDP-43^{Q3331K} mutation, can affect its binding to other proteins. The TDP-43^{Q331K} mutation has been suggested to result in a new, potential site of acetylation. In a study by Johnson *et al*, expression of this mutation in yeast cells results in increased aggregation propensity and cytotoxicity[154]. The authors of this study found the Q331K mutation to be the most cytotoxic compared to other TDP-43 variants and have noted that this form of the protein results in "more readily accessed threadlike aggregated forms", highlighting a possible avenue for its increased aggregation propensity.

Furthermore, the change from a charge-neutral residue like glutamine (Q) to a positively-charged amino acid like lysine (K) in the Q-to-K mutation could result in structural changes that affect the protein's stability, aggregation propensity, function and ability to form interactions, oligomerize or form stress granules[154]. Given that the CTD of the protein is unstructured, this single residue change could affect the stability of this region or result in the alteration of protein binding sites or pockets such that normal interactions can be hindered, while abnormal interactions be promoted, providing support for the toxic-gain-of-function mechanism.

Limitations

Our knock-in cell line is limited due to the cancerous nature of the HEK293T cell strain. HEK293T cells are neuroendocrine cells of fetal origin and utilised in many *in vitro* experiments due to ease of culture and transfection. Despite originating from adrenal tissue, these cells have been characterized to show that their genome and transcriptome is largely similar to that of neurons[155,156]. Additionally, ALS is a disease that primarily affects motor neurons and studies have observed altered axonal outgrowth and transport in neurons associated with TDP-43 pathology or other ALS-linked phenotypes; it is likely that we would not have identified proteins that are involved in these processes. Therefore, it is imperative that the hits identified using this cellular model be validated and characterized in other *in vitro* and *in vivo* models such as motor neurons or brain or spinal cord samples from ALS-mouse models.

Another limitation of our GFP-knock-in cell lines is the potential interruption to protein structure, function, and stability caused by the addition of the 28 kDa GFP-tag. Given the similar molecular weights of the GFP and TDP-43 proteins, conjugation of these components together may have large effects on protein function by hindering protein binding or TDP-43's ability to form complexes with other hnRNPs or RNA binding proteins. Furthermore, it could also influence the nuclear-to-cytoplasmic localization of TDP-43 which has been suggested to occur through passive diffusion out of the nucleus. In our cell line, the GFP-tag is localized to the C-terminal end of the protein and is attached to its unstructured prion-like domain in to have minimal functional impact. Furthermore, the addition of the GFP tag simplifies and overcomes many of the technical challenges posed with immunoprecipitation and subsequent mass spectrometry, as protocols

for untagged proteins need to be altered to include antibody: bead conjugation step, which is known to lead to antibody contamination and poorer sample quality for mass spectrometry[157].

While the addition of the GFP-tag provides ease-of-use of these KI cells in many different biochemical applications, including immunoprecipitation and microscopy, some additional characterization studies may be needed to assess the structural stability of GFP-tagged TDP-43, and to see whether the addition of this tag significantly alters the endogenous, biological nature of this model.

Additionally, it is important to note that many of TDP-43's protein interactions may be transient in nature or may be as part of a larger complex; as such, would not have been identified in our screen as they are not strong, direct interactions.

Despite these limitations, the proteins identified in our mass spectrometry dataset overlap in their function with many of TDP-43's known functions, both within the nucleus and cytoplasm. The use of our model is supported by the fact that several of our hits have been identified in other TDP-43 proteomics studies or genetic screens using both *in vitro* and *in vivo* models or linked to ALS through genetic data. TDP-43 is highly conserved, ubiquitously expression, and essential in all cell types; therefore, it is likely that many of its functions, or alterations to its pathways upon introduction of a mutation, will be retained in our cellular model. Our TDP-43-GFP KI cell lines have been generated using the HEK293T strain, which is known to be easy to genetically modify, relatively resistant to the stresses associated with cell culture applications and has a short growth/replication cycle—allowing us to easily expand our collection of cell lines to encompass other *TARDBP* mutations or to generate *wildtype* and mutant GFP-KIs for other ALS-linked genes, or to study other diseases. Interestingly, some of our top validated hits include genes, *DDX39b, FRG1*, and *PABPC1*, which are implicated in neuromuscular disorders such as multiple sclerosis, muscular dystrophy, facioscapulohumeral muscular dystrophy 1, spinal muscular atrophy and spinocerebellar ataxia. Given that these diseases share some muscle atrophy or degeneration phenotypes, further study of these hits may reveal proteins or pathways that contribute to these phenotypes.

Furthermore, the GFP-KI allows for the use of these cells in many cellular and biochemical applications including microscopy, immunoprecipitation, and western blots, among others, with the addition of the GFP-

tag allowing researchers to overcome technical challenges such as protein:antibody conjugation, staining and reagent availability, among others associated with these techniques. As such, this cellular tool can easily be used to perform experiments in a high-throughput manner.

Future directions

In our cells, the Q331K mutation does not show major loss of function or mislocalization phenotypes. In addition to this, there are no gross changes to cell morphology or the formation of any cytoplasmic TDP-43 aggregates, which are pathological phenotypes observed in ALS patient cells. This could be attributed to the fact that there may be cellular mechanisms to compensate for any losses in function, or that the pathways affected in our cellular model have redundant functions.

Despite these relatively small effects to TDP-43 function and localization, our mass spectrometry data shows that there is a bias in protein-protein interactions, favouring TDP-43^{WT}-GFP. This observation supports the idea that the TDP-43 interactome is altered in the context of ALS. Further characterization studies to assess TDP-43 loss-of-function, mislocalization and other aspects of TDP-43 pathology should be implemented in other models of TDP-43 dysfunction or in ALS-relevant contexts such as brain, spinal cord or motor neurons from ALS-relevant mouse models or patient derived cells or tissue samples. Another limitation of our approach is that we have likely missed capturing many of TDP-43's weaker or transient protein interactions, however, this can be overcome with the use of proximity-based labelling approaches such as APEX or BioID, which allow for biotin-labelling of a protein of interest such as TDP-43, and its neighbouring interactors[158]. These proteins can be purified from cellular lysates and identified using biotin affinity capture (streptavidin-immunoprecipitation) and subsequent LC-MS/MS.

TDP-43 is an RNA binding protein and many of its functions are mediated through the formation of RNA:protein complexes. Additionally, alterations in the ALS transcriptome have been observed through RNA-seq studies, and given these findings, identifying whether these altered protein-protein interactions are RNA-dependant or independent can provide some key knowledge in understanding the mechanisms underlying ALS.

Finally, the findings of this study can have broader implications for other TDP-43 proteinopathies such as the subsets of frontotemporal dementias patients with FTD-U and FTD-NMD, as well as disorders which similarly to ALS, result in cytoplasmic TDP-43 inclusions and share symptoms such as impairments to motor behaviour.

9. CONCLUSION

In this thesis, we have feasibly generated endogenously expressing TDP-43-GFP-knock-in *wildtype* and ALS- causing mutant cell lines. We have characterized them to show that the TDP-43^{Q331K}-GFP mutation, when expressed in cells shows loss of function, loss of protein-protein interactions and TDP-43 cytosolic mislocalization. We observed an overall loss of interactors in these cells, particularly of nuclear RNA binding proteins involved in functions such as splicing, mRNA processing and nuclear export. Of these proteins, we shortlisted 14 hits for validation and found strong, robust interaction of DDX39b, ELAVL1, PABPC1 and HNRNPC with wild-type TDP-43. This robust interaction was significantly reduced with the TDP-43^{Q331K} mutation for DDX39b and HNRNPC. The removal of an outlier in the PABPC1 GFP-IPs also suggests trending decreases in the degree of interaction with TDP-43^{Q331K}-GFP, while our results were inconclusive in the context of PABPC1. We have also generated overexpression and knockdown constructs to further interrogate these top hits and utilised these tools to assess TDP-43 localization using fluorescence microscopy, observing an increase in TDP-43 nuclear localization upon hit knockdown. Using these same knockdown and overexpression constructs, we do not observe significant TDP-43 loss of function using a previously-established RT-PCR based splicing assay.

Ultimately, the goal of our research is to expand our knock-in cell lines to include other ALS-causing TDP-43 mutations, as well as to include mutations in other ALS-linked proteins to build an ALS-wide interactome map. By using a ‘library’ of these ALS-relevant cell lines, we can identify shared and altered pathways that may be driving neurodegeneration and pathology in this disease. The findings from this undertaking will allow us to pinpoint specific proteins or pathways that lie at the core of this disease—and can be targeted as a therapeutic approach.

10. REFERENCES

1. Suk.T, and Rousseaux. M. 2020. The role of TDP-43 mislocalization in amyotrophic lateral sclerosis. *Mol. Neurodegener.* 15. doi:10.1186/S13024-020-00397-1.
2. Tetsuaki, A., H. Masato, A. Haruhiko *et al*. 2006. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* 351:602–611. doi:10.1016/J.BBRC.2006.10.093.
3. Sephton, C.F., B. Cenik, B.K. Cenik, J. Herz, and G. Yu. 2012. TDP-43 in CNS development and function: clues to TDP-43-associated neurodegeneration. *Biol. Chem.* 393:589. doi:10.1515/HSZ-2012-0115.
4. Sreedharan, J., I.P. Blair, V.B. Tripathi *et al*. 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. 319:1668–1672. doi:10.1126/SCIENCE.1154584.
5. Zarei, S., K. Carr, L. Reiley, K. Diaz, O. Guerra, P.F. Altamirano, W. Pagani, D. Lodin, G. Orozco, and A. Chinea. 2015. A comprehensive review of amyotrophic lateral sclerosis. *Surg. Neurol. Int.* 6. doi:10.4103/2152-7806.169561.
6. Kumar, D.R., F. Aslinia, S.H. Yale, and J.J. Mazza. 2011. Jean-Martin Charcot: The Father of Neurology. *Clin. Med. Res.* 9:46. doi:10.3121/CMR.2009.883.
7. Turner, M.R., and M. Swash. 2015. The expanding syndrome of amyotrophic lateral sclerosis: a clinical and molecular odyssey. *J. Neurol. Neurosurg. Psychiatry*. 86:667–673. doi:10.1136/JNNP-2014-308946.
8. Hulisz, D. 2018. Amyotrophic lateral sclerosis: disease state overview - PubMed. *Am J Manag Care.* 24:S320-S326.
9. Goetz, C.G. 2009. Chapter 15 Jean-Martin Charcot and the anatomo-clinical method of neurology. *Handb. Clin. Neurol.* 95:203–212. doi:10.1016/S0072-9752(08)02115-5.
10. Office of Neuroscience Communications and Engagement
National Institute of Neurological Disorders and Stroke. 2013. Amyotrophic Lateral Sclerosis (ALS) Fact Sheet | National Institute of Neurological Disorders and Stroke. *NIH Natl. Inst. Neurol. Disord. Stroke*.
11. Amyotrophic lateral sclerosis: MedlinePlus Genetics. *U.S Natl. Libr. Med., https://medlineplus.gov/genetics/condition/amyotrophic-lateral-sclerosis/#synonyms.*
12. Genetic Testing for ALS | The ALS Association. 2021. *ALS.ORG*. https://www.als.org/understanding-als/who-gets-als/genetic-testing
13. Wijesekera, L.C., and P.N. Leigh. 2009. Amyotrophic lateral sclerosis. *Orphanet J. Rare Dis.* 4:3. doi:10.1186/1750-1172-4-3.
14. Story: Limb and Bulbar Onset ALS | ALS Therapy Development Institute. 2021. *ALS Ther. Dev. Inst. https://www.als.net/news/science-sunday-limb-and-bulbar-onset-als/*
15. Finger, E.C. 2016. Frontotemporal Dementias. *Contin. Lifelong Learn. Neurol.* 22:464. doi:10.1212/CON.0000000000000300.
16. Bäumer, D., K. Talbot, and M.R. Turner. 2014. Advances in motor neurone disease. *J. R. Soc. Med.* 107:14. doi:10.1177/0141076813511451.
17. Desai, P., and G. Criner. 2017. Neuromuscular Disorders Affecting the Thorax: Amyotrophic Lateral Sclerosis - Pulmonology Advisor. *Pulm. Med.* .
18. Deng, Z., P. Sheehan, S. Chen, and Z. Yue. 2017. Is amyotrophic lateral sclerosis/frontotemporal dementia an autophagy disease? *Mol. Neurodegener.* 12. doi:10.1186/S13024-017-0232-6.
19. Masrori, P., and P. Van Damme. 2020. Amyotrophic lateral sclerosis: a clinical review. *Eur. J. Neurol.* 27:1918–1929. doi:10.1111/ENE.14393.
20. Xu, L., T. Liu, L. Liu *et al*. 2020. Global variation in prevalence and incidence of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J. Neurol.* 267:944–953. doi:10.1007/S00415-019-09652-Y.
21. About ALS. 2018. *ALS Soc. Canada*. https://www.als.ca/what-is-als/about-als/

22. Longinetti, E., and F. Fang. 2019. Epidemiology of amyotrophic lateral sclerosis: an update of recent literature. *Curr. Opin. Neurol.* 32:771. doi:10.1097/WCO.0000000000000730.
23. Ingre, C., P.M. Roos, F. Piehl, F. Kamel, and F. Fang. 2015. Risk factors for amyotrophic lateral sclerosis. *Clin. Epidemiol.* 7:181. doi:10.2147/CLEP.S37505.
24. Hergesheimer, R.C., A.A. Chami, D.R. de Assis, P. Vourc'h, C.R. Andres, P. Corcia, D. Lanznaster, and H. Blasco. 2019. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain*. 142:1176. doi:10.1093/BRAIN/AWZ078.
25. Cappella, M., P.-F. Pradat, G. Querin, and M.G. Biferi. 2021. Beyond the Traditional Clinical Trials for Amyotrophic Lateral Sclerosis and The Future Impact of Gene Therapy. *J. Neuromuscul. Dis.* 8:25. doi:10.3233/JND-200531.
26. Ling, J.P., O. Pletnikova, J.C. Troncoso, and P.C. Wong. 2015. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 349:650. doi:10.1126/SCIENCE.AAB0983.
27. Renton, A.E., A. Chiò, and B.J. Traynor. 2014. State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* 17:17. doi:10.1038/NN.3584.
28. Ferraiuolo, L., J. Kirby, A.J. Grierson, M. Sendtner, and P.J. Shaw. 2011. Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nat. Rev. Neurol. 2011 711*. 7:616–630. doi:10.1038/nrneurol.2011.152.
29. Chhangani, D., A. Martín-Peña, and D.E. Rincon-Limas. 2021. Molecular, functional, and pathological aspects of TDP-43 fragmentation. *iScience*. 24. doi:10.1016/J.ISCI.2021.102459.
30. Neumann, M., L. Kwong, E. Lee *et al*. 2009. Phosphorylation of S409/410 of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. *Acta Neuropathol.* 117:137–149. doi:10.1007/S00401-008-0477-9.
31. Al-Sarraj, S., A. King, C. Troakes *et al*. 2011. p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. *Acta Neuropathol.* 122:691–702. doi:10.1007/S00401-011-0911-2.
32. Torres, P., O. Ramírez-Núñez, R. Romero-Guevara, G. Barés, A.B. Granado-Serrano, V. Ayala, J. Boada, L. Fontdevila, M. Povedano, D. Sanchís, R. Pamplona, I. Ferrer, and M. Portero-Otín. 2018. Cryptic exon splicing function of TARDBP interacts with autophagy in nervous tissue. *Autophagy*. 14:1398. doi:10.1080/15548627.2018.1474311.
33. Ash, P.E.A., E.A. Stanford, A. Al Abdulatif, A. Ramirez-Cardenas, H.I. Ballance, S. Boudeau, A. Jeh, J.M. Murithi, Y. Tripodis, G.J. Murphy, D.H. Sherr, and B. Wolozin. 2017. Dioxins and related environmental contaminants increase TDP-43 levels. *Mol. Neurodegener. 2017 121*. 12:1–14. doi:10.1186/S13024-017-0177-9.
34. Trojsi, F., M.R. Monsurrò, and G. Tedeschi. 2013. Exposure to Environmental Toxicants and Pathogenesis of Amyotrophic Lateral Sclerosis: State of the Art and Research Perspectives. *Int. J. Mol. Sci.* 14:15286. doi:10.3390/IJMS140815286.
35. Sienko. D, Davis. J, Taylor. J, and Brooks. B. 1990. Amyotrophic lateral sclerosis. A case-control study following detection of a cluster in a small Wisconsin community. *Arch. Neurol.* 47:38–41. doi:10.1001/ARCHNEUR.1990.00530010046017.
36. Barberm TE. 1978. Inorganic mercury intoxication reminiscent of amyotrophic lateral sclerosis. *J. Occup. Med.* 20:667–669. PMID: 722351
37. Vinceti, M., G. Nacci, E. Rocchi *et al*. 2000. Mortality in a population with long-term exposure to inorganic selenium via drinking water. *J. Clin. Epidemiol.* 53:1062–1068. doi:10.1016/S0895-4356(00)00233-X.
38. Johnson, F.O., Y. Yuan, R.K. Hajela, A. Chitrakar, D.M. Parsell, and W.D. Atchison. 2011. Exposure to an Environmental Neurotoxicant Hastens the Onset of Amyotrophic Lateral Sclerosis-Like Phenotype in Human Cu2+/Zn2+ Superoxide Dismutase 1 G93A Mice: Glutamate-Mediated Excitotoxicity. *J. Pharmacol. Exp. Ther.* 338:518. doi:10.1124/JPET.110.174466.

39. Davidson-York, D., F. Galey, P. Blanchard, and I. Gardner. 1999. Selenium elimination in pigs after an outbreak of selenium toxicosis. *J. Vet. Diagn. Invest.* 11:352–357. doi:10.1177/104063879901100410.
40. Casteignau, A., A. Fontán, A. Morillo, J. Oliveros, and J. Segalés. 2006. Clinical, pathological and toxicological findings of a iatrogenic selenium toxicosis case in feeder pigs. *J. Vet. Med. A. Physiol. Pathol. Clin. Med.* 53:323–326. doi:10.1111/J.1439-0442.2006.00830.X.
41. Barbeito, A.G., L. Martinez-Palma, M.R. Vargas, M. Pehar, N. Mañay, J.S. Beckman, L. Barbeito, and P. Cassina. 2010. Lead exposure stimulates VEGF expression in the spinal cord and extends survival in a mouse model of ALS. *Neurobiol. Dis.* 37:574. doi:10.1016/J.NBD.2009.11.007.
42. Dickerson, A.S., J. Hansen, A.J. Specht, O. Gredal, and M.G. Weisskopf. 2019. Population-based study of amyotrophic lateral sclerosis and occupational lead exposure in Denmark. *Occup. Environ. Med.* 76:208–214. doi:10.1136/OEMED-2018-105469.
43. Okamoto, K., T. Kihira, G. Kobashi, M. Washio, S. Sasaki, T. Yokoyama, Y. Miyake, N. Sakamoto, Y. Inaba, and M. Nagai. 2009. Fruit and Vegetable Intake and Risk of Amyotrophic Lateral Sclerosis in Japan. *Neuroepidemiology*. 32:251–256. doi:10.1159/000201563.
44. Molina, J., F. de Bustos, F. Jiménez-Jiménez *et al*. 1999. Serum levels of beta-carotene, alpha-carotene, and vitamin A in patients with amyotrophic lateral sclerosis. *Acta Neurol. Scand.* 99:315–317. doi:10.1111/J.1600-0404.1999.TB00682.X.
45. Graf, M., D. Ecker, R. Horowski *et al*. 2005. High dose vitamin E therapy in amyotrophic lateral sclerosis as add-on therapy to riluzole: results of a placebo-controlled double-blind study. *J. Neural Transm.* 112:649–660. doi:10.1007/S00702-004-0220-1.
46. Nieves, J.W., C. Gennings, P. Factor-Litvak, J. Hupf, J. Singleton, V. Sharf, B. Oskarsson, J.A.M.F. Filho, E.J. Sorenson, E. D'Amico, R. Goetz, and H. Mitsumoto. 2016. Association Between Dietary Intake and Function in Amyotrophic Lateral Sclerosis. *JAMA Neurol.* 73:1425. doi:10.1001/JAMANEUROL.2016.3401.
47. Carrera-Juliá, S., M.L. Moreno, C. Barrios, J.E. de la R. Ortí, and E. Drehmer. 2020. Antioxidant Alternatives in the Treatment of Amyotrophic Lateral Sclerosis: A Comprehensive Review. *Front. Physiol.* 11:63. doi:10.3389/FPHYS.2020.00063.
48. Cruz, M.P. 2018. Edaravone (Radicava): A Novel Neuroprotective Agent for the Treatment of Amyotrophic Lateral Sclerosis. *Pharm. Ther.* 43:25.
49. Bozzoni, V., O. Pansarasa, L. Diamanti, G. Nosari, C. Cereda, and M. Ceroni. 2016. Amyotrophic lateral sclerosis and environmental factors. *Funct. Neurol.* 31:7. doi:10.11138/FNEUR/2016.31.1.007.
50. Lehman, E.J., M.J. Hein, S.L. Baron, and C.M. Gersic. 2012. Neurodegenerative causes of death among retired National Football League players. *Neurology*. 79:1970. doi:10.1212/WNL.0B013E31826DAF50.
51. Chiò, A., G. Benzi, M. Dossena *et al*. 2005. Severely increased risk of amyotrophic lateral sclerosis among Italian professional football players. *Brain*. 128:472–476. doi:10.1093/BRAIN/AWH373.
52. Harwood, C.A., C.J. McDermott, and P.J. Shaw. 2009. Physical activity as an exogenous risk factor in motor neuron disease (MND): A review of the evidence. *http://dx.doi.org/10.1080/17482960802549739*. 10:191–204. doi:10.1080/17482960802549739.
53. Ingre, C., P.M. Roos, F. Piehl, F. Kamel, and F. Fang. 2015. Risk factors for amyotrophic lateral sclerosis. *Clin. Epidemiol.* 7:181. doi:10.2147/CLEP.S37505.
54. Hu, M., and N. Robertson. 2020. Physical activity as a risk factor for amyotrophic lateral sclerosis-findings from three large European cohorts. *J. Neurol. 2020 2677*. 267:2173–2175. doi:10.1007/S00415-020-09995-X.
55. Trojsi, F., M.R. Monsurrò, and G. Tedeschi. 2013. Exposure to Environmental Toxicants and Pathogenesis of Amyotrophic Lateral Sclerosis: State of the Art and Research Perspectives. *Int. J. Mol. Sci.* 14:15286. doi:10.3390/IJMS140815286.

56. Håkansson, N., P. Gustavsson, C. Johansen, and B. Floderus. 2003. Neurodegenerative diseases in welders and other workers exposed to high levels of magnetic fields. *Epidemiology*. 14:420–426. doi:10.1097/01.EDE.0000078446.76859.C9.
57. Filippini, T., E.E. Hatch, and M. Vinceti. 2021. Residential exposure to electromagnetic fields and risk of amyotrophic lateral sclerosis: a dose–response meta-analysis. *Sci. Reports 2021 111*. 11:1–9. doi:10.1038/s41598-021-91349-2.
58. Zhou, H., G. Chen, C. Chen, Y. Yu, and Z. Xu. 2012. Association between Extremely Low-Frequency Electromagnetic Fields Occupations and Amyotrophic Lateral Sclerosis: A Meta-Analysis. *PLoS One*. 7:e48354. doi:10.1371/JOURNAL.PONE.0048354.
59. Gordon, P.H. 2013. Amyotrophic Lateral Sclerosis: An update for 2013 Clinical Features, Pathophysiology, Management and Therapeutic Trials. *Aging Dis.* 4:295. doi:10.14336/AD.2013.0400295.
60. Hinchcliffe, M., and A. Smith. 2017. Riluzole: real-world evidence supports significant extension of median survival times in patients with amyotrophic lateral sclerosis. *Degener. Neurol. Neuromuscul. Dis.* 7:61. doi:10.2147/DNND.S135748.
61. FDA-Approved Drugs | The ALS Association. 2021. *ALS Assoc. https://www.als.org/navigating-als/living-with-als/fda-approved-drugs*
62. Approved Treatments - ALS News Today. 2020. *ALS News Today. https://alsnewstoday.com/approved-treatments/*
63. Cruz, M.P. 2018. Edaravone (Radicava): A Novel Neuroprotective Agent for the Treatment of Amyotrophic Lateral Sclerosis. *Pharm. Ther.* 43:25.
64. Miller, R.G., J.D. Mitchell, and D.H. Moore. 2012. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst. Rev.* doi:10.1002/14651858.CD001447.PUB3.
65. Lacomblez, L., G. Bensimon, P. Leigh *et al*. 1996. Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II. *Lancet (London, England)*. 347:1425–1431. doi:10.1016/S0140-6736(96)91680-3.
66. Hinchcliffe, M., and A. Smith. 2017. Riluzole: real-world evidence supports significant extension of median survival times in patients with amyotrophic lateral sclerosis. *Degener. Neurol. Neuromuscul. Dis.* 7:61. doi:10.2147/DNND.S135748.
67. Cruz, M.P. 2013. Nuedexta for the Treatment Of Pseudobulbar Affect: A Condition of Involuntary Crying or Laughing. *Pharm. Ther.* 38:325.
68. Young, J.J., M. Lavakumar, D. Tampi, S. Balachandran, and R.R. Tampi. 2018. Frontotemporal dementia: latest evidence and clinical implications. *Ther. Adv. Psychopharmacol.* 8:33. doi:10.1177/2045125317739818.
69. Mackenzie, I.R.A., and R. Rademakers. 2008. The role of TDP-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Curr. Opin. Neurol.* 21:693. doi:10.1097/WCO.0B013E3283168D1D.
70. Olney, N.T., S. Spina, and B.L. Miller. 2017. Frontotemporal Dementia. *Neurol. Clin.* 35:339. doi:10.1016/J.NCL.2017.01.008.
71. Birsa, N., M. Bentham, and P. Fratta. 2020. Cytoplasmic functions of TDP-43 and FUS and their role in ALS. *Semin. Cell Dev. Biol.* 99:193–201. doi:10.1016/J.SEMCDB.2019.05.023.
72. Huang, C., S. Yan, and Z. Zhang. 2020. Maintaining the balance of TDP-43, mitochondria, and autophagy: a promising therapeutic strategy for neurodegenerative diseases. *Transl. Neurodegener. 2020 91*. 9:1–16. doi:10.1186/S40035-020-00219-W.
73. Ederle, H., and D. Dormann. 2017. TDP-43 and FUS en route from the nucleus to the cytoplasm. *FEBS Lett.* 591:1489–1507. doi:10.1002/1873-3468.12646.
74. Polymenidou, M., C. Lagier-Tourenne, K.R. Hutt, S.C. Huelga, J. Moran, T.Y. Liang, S.-C. Ling, E. Sun, E. Wancewicz, C. Mazur, H. Kordasiewicz, Y. Sedaghat, J. Paul Donohue, L. Shiue, C. Frank Bennett, G.W. Yeo, D.W. Cleveland, and N. Neurosci. 2011. Long pre-mRNA depletion

and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43 HHS Public Access Author manuscript. *Nat Neurosci*. 14:459–468. doi:10.1038/nn.2779.
75. Polymenidou, M., C. Lagier-Tourenne, K. Hutt *et al*. 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* 14:459–468. doi:10.1038/NN.2779.
76. Ayala, Y.M., L. De Conti, S.E. Avendaño-Vázquez, A. Dhir, M. Romano, A. D'Ambrogio, J. Tollervey, J. Ule, M. Baralle, E. Buratti, and F.E. Baralle. 2011. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* 30:277. doi:10.1038/EMBOJ.2010.310.
77. Ling, J.P., O. Pletnikova, J.C. Troncoso, and P.C. Wong. 2015. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 349:650. doi:10.1126/SCIENCE.AAB0983.
78. Nonaka, T., M. Masuda-Suzukake, T. Arai, Y. Hasegawa, H. Akatsu, T. Obi, M. Yoshida, S. Murayama, D.M.A. Mann, H. Akiyama, and M. Hasegawa. 2013. Prion-like Properties of Pathological TDP-43 Aggregates from Diseased Brains. *Cell Rep.* 4:124–134. doi:10.1016/J.CELREP.2013.06.007.
79. Xiao, S., T. Sanelli, H. Chiang, Y. Sun, A. Chakrabartty, J. Keith, E. Rogaeva, L. Zinman, and J. Robertson. 2015. Low molecular weight species of TDP-43 generated by abnormal splicing form inclusions in amyotrophic lateral sclerosis and result in motor neuron death. *Acta Neuropathol.* 130:49. doi:10.1007/S00401-015-1412-5.
80. Arai, T., H. M, T. Nonoka *et al*. 2010. Phosphorylated and cleaved TDP-43 in ALS, FTLD and other neurodegenerative disorders and in cellular models of TDP-43 proteinopathy. *Neuropathology*. 30:170–181. doi:10.1111/J.1440-1789.2009.01089.X.
81. Hergesheimer, R.C., A.A. Chami, D.R. de Assis, P. Vourc'h, C.R. Andres, P. Corcia, D. Lanznaster, and H. Blasco. 2019. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain*. 142:1176. doi:10.1093/BRAIN/AWZ078.
82. Prasad, A., V. Bharathi, V. Sivalingam, A. Girdhar, and B.K. Patel. 2019. Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.* 0:25. doi:10.3389/FNMOL.2019.00025.
83. François-Moutal, L., S. Perez-Miller, D.D. Scott, V.G. Miranda, N. Mollasalehi, and M. Khanna. 2019. Structural Insights Into TDP-43 and Effects of Post-translational Modifications. *Front. Mol. Neurosci.* 12. doi:10.3389/FNMOL.2019.00301.
84. Archbold, H.C., K.L. Jackson, A. Arora, K. Weskamp, E.M.-H. Tank, X. Li, R. Miguez, R.D. Dayton, S. Tamir, R.L. Klein, and S.J. Barmada. 2018. TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci. Reports 2018 81*. 8:1–18. doi:10.1038/s41598-018-22858-w.
85. Giannini, M., A. Bayona-Feliu, D. Sproviero, S.I. Barroso, C. Cereda, and A. Aguilera. 2020. TDP-43 mutations link Amyotrophic Lateral Sclerosis with R-loop homeostasis and R loop-mediated DNA damage. *PLOS Genet.* 16:e1009260. doi:10.1371/JOURNAL.PGEN.1009260.
86. Valdmanis, P.N., H. Daoud, P.A. Dion, and G.A. Rouleau. 2009. Recent Advances in the Genetics of Amyotrophic Lateral Sclerosis.
87. Gendron, T.F., R. Rademakers, and L. Petrucelli. 2013. TARDBP mutation analysis in TDP-43 proteinopathies and deciphering the toxicity of mutant TDP-43. *J. Alzheimers. Dis.* 33:S35. doi:10.3233/JAD-2012-129036.
88. Lattante, S., G.A. Rouleau, and E. Kabashi. 2013. TARDBP and FUS Mutations Associated with Amyotrophic Lateral Sclerosis: Summary and Update. *Hum. Mutat.* 34:812–826. doi:10.1002/HUMU.22319.
89. Johnson, B., D. Snead, and J. Lee *at al*. 2009. TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* 284:20329–20339. doi:10.1074/JBC.M109.010264.
90. Wobst, H.J., S.S. Wesolowski, J. Chadchankar, L. Delsing, S. Jacobsen, J. Mukherjee, T.Z. Deeb, J. Dunlop, N.J. Brandon, and S.J. Moss. 2017. Cytoplasmic Relocalization of TAR DNA-Binding

Protein 43 Is Not Sufficient to Reproduce Cellular Pathologies Associated with ALS In vitro. *Front. Mol. Neurosci.* 10. doi:10.3389/FNMOL.2017.00046.

91. Sreedharan, J., I.P. Blair, V.B. Tripathi, X. Hu, C. Vance, B. Rogelj, S. Ackerley, J.C. Durnall, K.L. Williams, E. Buratti, F. Baralle, J. de Belleroche, J.D. Mitchell, P.N. Leigh, A. Al-Chalabi, C.C. Miller, G. Nicholson, and C.E. Shaw. 2008. TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science*. 319:1668. doi:10.1126/SCIENCE.1154584.
92. Mitchell, J.C., R. Constable, E. So, C. Vance, E. Scotter, L. Glover, T. Hortobagyi, E.S. Arnold, S.-C. Ling, M. McAlonis, S. Da Cruz, M. Polymenidou, L. Tessarolo, D.W. Cleveland, and C.E. Shaw. 2015. Wild type human TDP-43 potentiates ALS-linked mutant TDP-43 driven progressive motor and cortical neuron degeneration with pathological features of ALS. *Acta Neuropathol. Commun. 2015 31*. 3:1–16. doi:10.1186/S40478-015-0212-4.
93. Chand, K.K., K.M. Lee, J.D. Lee, H. Qiu, E.F. Willis, N.A. Lavidis, M.A. Hilliard, and P.G. Noakes. 2018. Defects in synaptic transmission at the neuromuscular junction precede motor deficits in a TDP-43Q331K transgenic mouse model of amyotrophic lateral sclerosis. *FASEB J.* 32:2676–2689. doi:10.1096/FJ.201700835R.
94. Liu, R., G. Yang, T. Nonaka, T. Arai, W. Jia, and M.S. Cynader. 2013. Reducing TDP-43 aggregation does not prevent its cytotoxicity. *Acta Neuropathol. Commun. 2013 11*. 1:1–11. doi:10.1186/2051-5960-1-49.
95. Tanji, K., H. Zhang, F. Mori *et al*. 2012. p62/sequestosome 1 binds to TDP-43 in brains with frontotemporal lobar degeneration with TDP-43 inclusions. *J. Neurosci. Res.* 90:2034–2042. doi:10.1002/JNR.23081.
96. Gal, J., A. Ström, R. Kilty, F. Zhang, and H. Zhu. 2007. p62 accumulates and enhances aggregate formation in model systems of familial amyotrophic lateral sclerosis. *J. Biol. Chem.* 282:11068–11077. doi:10.1074/JBC.M608787200.
97. Seibenhener, M.L., J.R. Babu, T. Geetha, H.C. Wong, N.R. Krishna, and M.W. Wooten. 2004. Sequestosome 1/p62 Is a Polyubiquitin Chain Binding Protein Involved in Ubiquitin Proteasome Degradation. *Mol. Cell. Biol.* 24:8055. doi:10.1128/MCB.24.18.8055-8068.2004.
98. Ayala, Y., S. Pantano, A. D'Ambrogio *et al*. 2005. Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol.* 348:575–588. doi:10.1016/J.JMB.2005.02.038.
99. Yang, C., H. Wang, T. Qiao, B. Yang, L. Aliaga, L. Qiu, W. Tan, J. Salameh, D.M. McKenna-Yasek, T. Smith, L. Peng, M.J. Moore, R.H. Brown, H. Cai, and Z. Xu. 2014. Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* 111:E1121–E1129. doi:10.1073/PNAS.1322641111.
100. Zhang, Y., T. Caulfield, Y. Xu *et al*. 2013. The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* 22:3112–3122. doi:10.1093/HMG/DDT166.
101. Walker, A., K. Spiller, G. Ge *et al*. 2015. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. *Acta Neuropathol.* 130:643–660. doi:10.1007/S00401-015-1460-X.
102. Vanden Broeck, L., P. Callaerts, and B. Dermaut. 2014. TDP-43-mediated neurodegeneration: towards a loss-of-function hypothesis? *Trends Mol. Med.* 20:66–71. doi:10.1016/j.molmed.2013.11.003.
103. Lin, M.-J., C.-W. Cheng, and C.-K.J. Shen. 2011. Neuronal Function and Dysfunction of Drosophila dTDP. *PLoS One*. 6:e20371. doi:10.1371/JOURNAL.PONE.0020371.
104. Diaper, D., Y. Adachi, B. Sutcliffe *et al*. 2013. Loss and gain of Drosophila TDP-43 impair synaptic efficacy and motor control leading to age-related neurodegeneration by loss-of-function phenotypes. *Hum. Mol. Genet.* 22:1539–1557. doi:10.1093/HMG/DDT005.
105. Iguchi, Y., M. Katsuno, J. Niwa, S. Takagi, S. Ishigaki, K. Ikenaka, K. Kawai, H. Watanabe, K. Yamanaka, R. Takahashi, H. Misawa, S. Sasaki, F. Tanaka, and G. Sobue. 2013. Loss of TDP-43

causes age-dependent progressive motor neuron degeneration. *Brain*. 136:1371–1382. doi:10.1093/BRAIN/AWT029.
106. Kabashi, E., L. Lin, M.L. Tradewell, P.A. Dion, V. Bercier, P. Bourgouin, D. Rochefort, S. Bel Hadj, H.D. Durham, C. Vande Velde, G.A. Rouleau, and P. Drapeau. 2010. Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum. Mol. Genet.* 19:671–683. doi:10.1093/HMG/DDP534.
107. Hanson, K.A., S.H. Kim, D.A. Wassarman, and R.S. Tibbetts. 2010. Ubiquilin Modifies TDP-43 Toxicity in a Drosophila Model of Amyotrophic Lateral Sclerosis (ALS). *J. Biol. Chem.* 285:11068. doi:10.1074/JBC.C109.078527.
108. White, M., E. Kim, A. Duffy *et al*. 2018. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat. Neurosci.* 21:552–563. doi:10.1038/S41593-018-0113-5.
109. Freibaum, B.D., R. Chitta, A.A. High, and J.P. Taylor. 2010. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J. Proteome Res.* 9:1104. doi:10.1021/PR901076Y.
110. Ling, S.-C., C.P. Albuquerque, J.S. Han, C. Lagier-Tourenne, S. Tokunaga, H. Zhou, and D.W. Cleveland. ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. doi:10.1073/pnas.1008227107.
111. Sephton, C.F., C. Cenik, A. Kucukural, E.B. Dammer, B. Cenik, Y. Han, C.M. Dewey, F.P. Roth, J. Herz, J. Peng, M.J. Moore, and G. Yu. 2011. Identification of Neuronal RNA Targets of TDP-43-containing Ribonucleoprotein Complexes. *J. Biol. Chem.* 286:1204. doi:10.1074/JBC.M110.190884.
112. Colombrita, C., E. Onesto, F. Megiorni, A. Pizzuti, F.E. Baralle, E. Buratti, V. Silani, and A. Ratti. 2012. TDP-43 and FUS RNA-binding Proteins Bind Distinct Sets of Cytoplasmic Messenger RNAs and Differently Regulate Their Post-transcriptional Fate in Motoneuron-like Cells *. *J. Biol. Chem.* 287:15635–15647. doi:10.1074/JBC.M111.333450.
113. Kawaguchi, T., R. Rollins, M. Moinpour *et al*. 2020. Changes to the TDP-43 and FUS Interactomes Induced by DNA Damage. *J. Proteome Res.* 19:360–370. doi:10.1021/ACS.JPROTEOME.9B00575.
114. Feneberg, E., D. Gordon, A.G. Thompson, M.J. Finelli, R. Dafinca, A. Candalija, P.D. Charles, I. Mäger, M.J. Wood, R. Fischer, B.M. Kessler, E. Gray, M.R. Turner, and K. Talbot. 2020. An ALS-linked mutation in TDP-43 disrupts normal protein interactions in the motor neuron response to oxidative stress. *Neurobiol. Dis.* 144:105050. doi:10.1016/J.NBD.2020.105050.
115. Pédelacq, J.-D., S. Cabantous, T. Tran, T.C. Terwilliger, and G.S. Waldo. 2005. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol. 2005 241*. 24:79–88. doi:10.1038/nbt1172.
116. Shevchenko, A., H. Tomas, J. Havlis, J. Olsen, and M. Mann. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1:2856–2860. doi:10.1038/NPROT.2006.468.
117. R, S., and S. H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*. 22:1540–1542. doi:10.1093/BIOINFORMATICS/BTL117.
118. Choi, H., B. Larsen, Z. Lin *et al*. 2011. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat. Methods*. 8:70–73. doi:10.1038/NMETH.1541.
119. UniProt: the universal protein knowledgebase in 2021. 2021. *Nucleic Acids Res.* 49:D480–D489. doi:10.1093/NAR/GKAA1100.
120. Yang, X., J.S. Boehm, X. Yang, K. Salehi-Ashtiani, T. Hao, Y. Shen, R. Lubonja, S.R. Thomas, O. Alkan, T. Bhimdi, T.M. Green, C.M. Johannessen, S.J. Silver, C. Nguyen, R.R. Murray, H. Hieronymus, D. Balcha, C. Fan, C. Lin, L. Ghamsari, M. Vidal, W.C. Hahn, D.E. Hill, and D.E. Root. 2011. A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods 2011 88*. 8:659–661. doi:10.1038/nmeth.1638.

121. McQuin, C., A. Goodman, V. Chernyshev, L. Kamentsky, B.A. Cimini, K.W. Karhohs, M. Doan, L. Ding, S.M. Rafelski, D. Thirstrup, W. Wiegraebe, S. Singh, T. Becker, J.C. Caicedo, and A.E. Carpenter. 2018. CellProfiler 3.0: Next-generation image processing for biology. *PLOS Biol.* 16:e2005970. doi:10.1371/JOURNAL.PBIO.2005970.
122. Costessi, L., F. Porro, A. Iaconcig, and A. Muro. 2014. TDP-43 regulates β-adducin (Add2) transcript stability. *RNA Biol.* 11:1280–1290. doi:10.1080/15476286.2014.996081.
123. Elden, A.C., H.-J. Kim, M.P. Hart, A.S. Chen-Plotkin, B.S. Johnson, X. Fang, M. Armakola, F. Geser, R. Greene, M.M. Lu, A. Padmanabhan, D. Clay, L. McCluskey, L. Elman, D. Juhr, P.J. Gruber, U. Rüb, G. Auburger, J.Q. Trojanowski, V.M.-Y. Lee, V.M. Van Deerlin, N.M. Bonini, and A.D. Gitler. 2010. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. 466:1069. doi:10.1038/NATURE09320.
124. Vogt, M.A., Z. Ehsaei, P. Knuckles, A. Higginbottom, M.S. Helmbrecht, T. Kunath, K. Eggan, L.A. Williams, P.J. Shaw, W. Wurst, T. Floss, A.B. Huber, and V. Taylor. 2018. TDP-43 induces p53-mediated cell death of cortical progenitors and immature neurons. *Sci. Rep.* 8:8097. doi:10.1038/S41598-018-26397-2.
125. Klim, J.R., L.A. Williams, F. Limone, I.G.S. Juan, B.N. Davis-Dusenbery, D.A. Mordes, A. Burberry, M.J. Steinbaugh, K.K. Gamage, R. Kirchner, R. Moccia, S.H. Cassel, K. Chen, B.J. Wainger, C.J. Woolf, and K. Eggan. 2019. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat. Neurosci. 2019 222*. 22:167–179. doi:10.1038/s41593-018-0300-4.
126. Theunissen, F., R.S. Anderton, F.L. Mastaglia, L.L. Flynn, S.J. Winter, I. James, R. Bedlack, S. Hodgetts, S. Fletcher, S.D. Wilton, N.G. Laing, M. MacShane, M. Needham, A. Saunders, A. Mackay-Sim, Z. Melamed, J. Ravits, D.W. Cleveland, and P.A. Akkari. 2021. Novel STMN2 Variant Linked to Amyotrophic Lateral Sclerosis Risk and Clinical Phenotype. *Front. Aging Neurosci.* 0:127. doi:10.3389/FNAGI.2021.658226.
127. Melamed, Z., J. Lopez-Erauskin, M.W. Baughn, O. Zhang, K. Drenner, Y. Sun, F. Freyermuth, M.A. McMahon, M.S. Beccari, J. Artates, T. Ohkubo, M. Rodriguez, N. Lin, D. Wu, C.F. Bennett, F. Rigo, S. Da Cruz, J. Ravits, C. Lagier-Tourenne, and D.W. Cleveland. 2019. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat. Neurosci.* 22:180. doi:10.1038/S41593-018-0293-Z.
128. Geuens, T., D. Bouhy, and V. Timmerman. 2016. The hnRNP family: insights into their role in health and disease. *Hum. Genet.* 135:851. doi:10.1007/S00439-016-1683-5.
129. Heath, C., N. Viphakone, and S. Wilson. 2016. The role of TREX in gene expression and disease. *Biochem. J.* 473:2911–2935. doi:10.1042/BCJ20160010.
130. Pühringer, T., U. Hohmann, L. Fin *et al*. 2020. Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *Elife*. 9:1–65. doi:10.7554/ELIFE.61503.
131. Chi, B., Q. Wang, G. Wu, M. Tan, L. Wang, M. Shi, X. Chang, and H. Cheng. 2013. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.* 41:1294–1306. doi:10.1093/NAR/GKS1188.
132. Couthouis, J., M. Hart, J. Shorter *et al*. 2011. A yeast functional screen predicts new candidate ALS disease genes. *Proc. Natl. Acad. Sci. U. S. A.* 108:20881–20890. doi:10.1073/PNAS.1109434108.
133. Berson, A., L.D. Goodman, A.N. Sartoris, C.G. Otte, J.A. Aykit, V.M.-Y. Lee, J.Q. Trojanowski, and N.M. Bonini. 2019. Drosophila Ref1 / ALYREF regulates transcription and toxicity associated with ALS/FTD disease etiologies. *Acta Neuropathol. Commun. 2019 71*. 7:1–10. doi:10.1186/S40478-019-0710-X.
134. Appocher, C., F. Mohagheghi, S. Cappelli, C. Stuani, M. Romano, F. Feiguin, and E. Buratti. 2017. Major hnRNP proteins act as general TDP-43 functional modifiers both in Drosophila and human neuronal cells. *Nucleic Acids Res.* 45:8026. doi:10.1093/NAR/GKX477.

135. Purice, M.D., and J.P. Taylor. 2018. Linking hnRNP Function to ALS and FTD Pathology. *Front. Neurosci.* 12:326. doi:10.3389/FNINS.2018.00326.
136. Galarza Munoz, G. Regulation of IL7R splicing by DDX39B and its role in Multiple Sclerosis.
137. Hanel, M., R. Wuebbles, and P. Jones. 2009. Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Dev. Dyn.* 238:1502–1512. doi:10.1002/DVDY.21830.
138. Donlin-Asp, P.G., C. Fallini, J. Campos, H.C. Phan, G.J. Bassell, and W. Rossoll. 2017. The Survival of Motor Neuron Protein Acts as a Molecular Chaperone for mRNP Assembly. *Cell Rep.* 18:1660–1673. doi:10.1016/j.celrep.2017.01.059.
139. Damrath, E., M. Heck, S. Gispert *et al*. 2012. ATXN2-CAG42 sequesters PABPC1 into insolubility and induces FBXW8 in cerebellum of old ataxic knock-in mice. *PLoS Genet.* 8. doi:10.1371/JOURNAL.PGEN.1002920.
140. De Conti, L., M. V. Akinyi, R. Mendoza-Maldonado, M. Romano, M. Baralle, and E. Buratti. 2015. TDP-43 affects splicing profiles and isoform production of genes involved in the apoptotic and mitotic cellular pathways. *Nucleic Acids Res.* 43:8990. doi:10.1093/NAR/GKV814.
141. Shiga, A., T. Ishihara, A. Miyashita, M. Kuwabara, T. Kato, N. Watanabe, A. Yamahira, C. Kondo, A. Yokoseki, M. Takahashi, R. Kuwano, A. Kakita, M. Nishizawa, H. Takahashi, and O. Onodera. 2012. Alteration of POLDIP3 Splicing Associated with Loss of Function of TDP-43 in Tissues Affected with ALS. *PLoS One*. 7:e43120. doi:10.1371/JOURNAL.PONE.0043120.
142. Dudman. J, and Qi. X. 2020. Stress Granule Dysregulation in Amyotrophic Lateral Sclerosis. *Front. Cell. Neurosci.* 14. doi:10.3389/FNCEL.2020.598517.
143. Bekenstein, U., and H. Soreq. 2013. Heterogeneous nuclear ribonucleoprotein A1 in health and neurodegenerative disease: from structural insights to post-transcriptional regulatory roles. *Mol. Cell. Neurosci.* 56:436–446. doi:10.1016/J.MCN.2012.12.002.
144. Ma, W.J., S. Cheng, C. Campbell, A. Wright, and H. Furneaux. 1996. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J. Biol. Chem.* 271:8144–8151. doi:10.1074/JBC.271.14.8144.
145. Couthouis, J., M.P. Hart, J. Shorter, M. DeJesus-Hernandez, R. Erion, R. Oristano, A.X. Liu, D. Ramos, N. Jethava, D. Hosangadi, J. Epstein, A. Chiang, Z. Diaz, T. Nakaya, F. Ibrahim, H.-J. Kim, J.A. Solski, K.L. Williams, J. Mojsilovic-Petrovic, C. Ingre, K. Boylan, N.R. Graff-Radford, D.W. Dickson, D. Clay-Falcone, L. Elman, L. McCluskey, R. Greene, R.G. Kalb, V.M.-Y. Lee, J.Q. Trojanowski, A. Ludolph, W. Robberecht, P.M. Andersen, G.A. Nicholson, I.P. Blair, O.D. King, N.M. Bonini, V. Van Deerlin, R. Rademakers, Z. Mourelatos, and A.D. Gitler. 2011. A yeast functional screen predicts new candidate ALS disease genes. *Proc. Natl. Acad. Sci.* 108:20881–20890. doi:10.1073/PNAS.1109434108.
146. Kim, H.-J., A.R. Raphael, E.S. LaDow, L. McGurk, R.A. Weber, J.Q. Trojanowski, V.M.-Y. Lee, S. Finkbeiner, A.D. Gitler, and N.M. Bonini. 2013. Therapeutic modulation of eIF2α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nat. Genet. 2013 462*. 46:152–160. doi:10.1038/ng.2853.
147. Feneberg, E., D. Gordon, A.G. Thompson, M.J. Finelli, R. Dafinca, A. Candalija, P.D. Charles, I. Mäger, M.J. Wood, R. Fischer, B.M. Kessler, E. Gray, M.R. Turner, and K. Talbot. 2020. An ALS-linked mutation in TDP-43 disrupts normal protein interactions in the motor neuron response to oxidative stress. *Neurobiol. Dis.* 144:105050. doi:10.1016/J.NBD.2020.105050.
148. Damrath, E., M. Heck, S. Gispert *et al*. 2012. ATXN2-CAG42 sequesters PABPC1 into insolubility and induces FBXW8 in cerebellum of old ataxic knock-in mice. *PLoS Genet.* 8. doi:10.1371/JOURNAL.PGEN.1002920.
149. Hirsch-Reinshagen, V., C. Pottier, A.M. Nicholson, M. Baker, G.-Y.R. Hsiung, C. Krieger, P. Sengdy, K.B. Boylan, D.W. Dickson, M. Mesulam, S. Weintraub, E. Bigio, L. Zinman, J. Keith, E. Rogaeva, S.A. Zivkovic, D. Lacomis, J.P. Taylor, R. Rademakers, and I.R.A. Mackenzie. 2017. Clinical and neuropathological features of ALS/FTD with TIA1 mutations. doi:10.14288/1.0368584.

150. Aulas, A., and C. Vande Velde. 2015. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? *Front. Cell. Neurosci.* 9. doi:10.3389/FNCEL.2015.00423.
151. Kim, W., D. Kim, and K. Lee. 2021. RNA-Binding Proteins and the Complex Pathophysiology of ALS. *Int. J. Mol. Sci.* 22:1–18. doi:10.3390/IJMS22052598.
152. Szymura, S.J., G.M. Bernal, L. Wu, Z. Zhang, C.D. Crawley, D.J. Voce, P.-A. Campbell, D.E. Ranoa, R.R. Weichselbaum, and B. Yamini. 2020. DDX39B interacts with the pattern recognition receptor pathway to inhibit NF-κB and sensitize to alkylating chemotherapy. *BMC Biol. 2020 181*. 18:1–17. doi:10.1186/S12915-020-0764-Z.
153. Boehringer, A., K. Garcia-Mansfield, G. Singh, N. Bakkar, P. Pirrotte, and R. Bowser. 2017. ALS Associated Mutations in Matrin 3 Alter Protein-Protein Interactions and Impede mRNA Nuclear Export. *Sci. Reports 2017 71*. 7:1–14. doi:10.1038/s41598-017-14924-6.
154. Johnson, B., D. Snead, and J. Lee. 2009. TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* 284:20329–20339. doi:10.1074/JBC.M109.010264.
155. Lin, Y.-C., M. Boone, L. Meuris, I. Lemmens, N. Van Roy, A. Soete, J. Reumers, M. Moisse, S. Plaisance, R. Drmanac, J. Chen, F. Speleman, D. Lambrechts, Y. Van de Peer, J. Tavernier, and N. Callewaert. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat. Commun. 2014 51*. 5:1–12. doi:10.1038/ncomms5767.
156. Shaw, G., S. Morse, M. Ararat, and F.L. Graham. 2002. Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *FASEB J.* 16:869–871. doi:10.1096/FJ.01-0995FJE.
157. Kaboord, B., S. Smith, B. Patel, and S. Meier. 2015. Enrichment of low-abundant protein targets by immunoprecipitation upstream of mass spectrometry. *Methods Mol. Biol.* 1295:135–151. doi:10.1007/978-1-4939-2550-6_12.
158. Trinkle-Mulcahy, L. 2019. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Research*. 8. doi:10.12688/F1000RESEARCH.16903.

11. SUPPLEMENTARY FIGURES

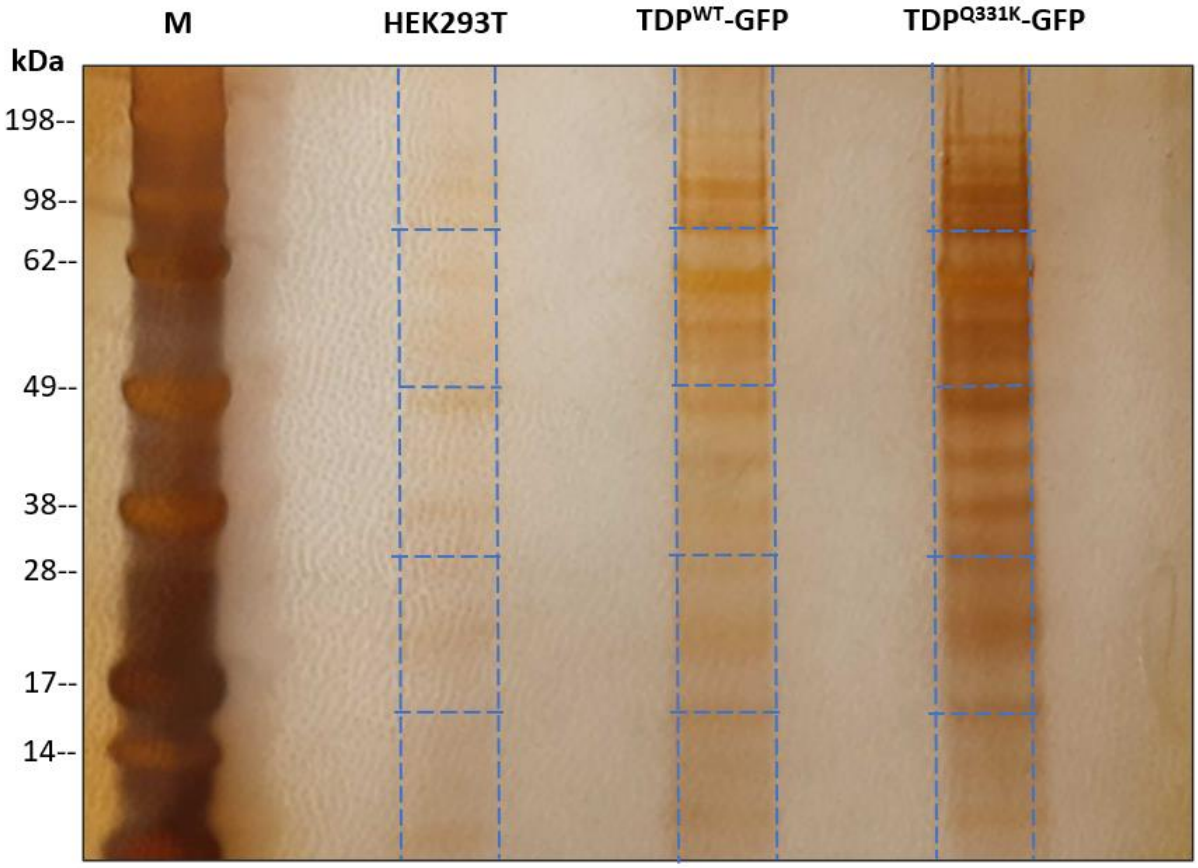


Figure S1: Silver stained SDSPAGE gel of GFP-IP experiment for analysis by mass spectrometry. HEK293T, TDP-43WT-GFP and TDP-43Q331K-GFP cell lysates were used to perform GFP-IP. IP samples were gel separated using SDSPAGE and subsequent silver-staining was performed. Gel bands were excised as indicated by blue dotted lines.

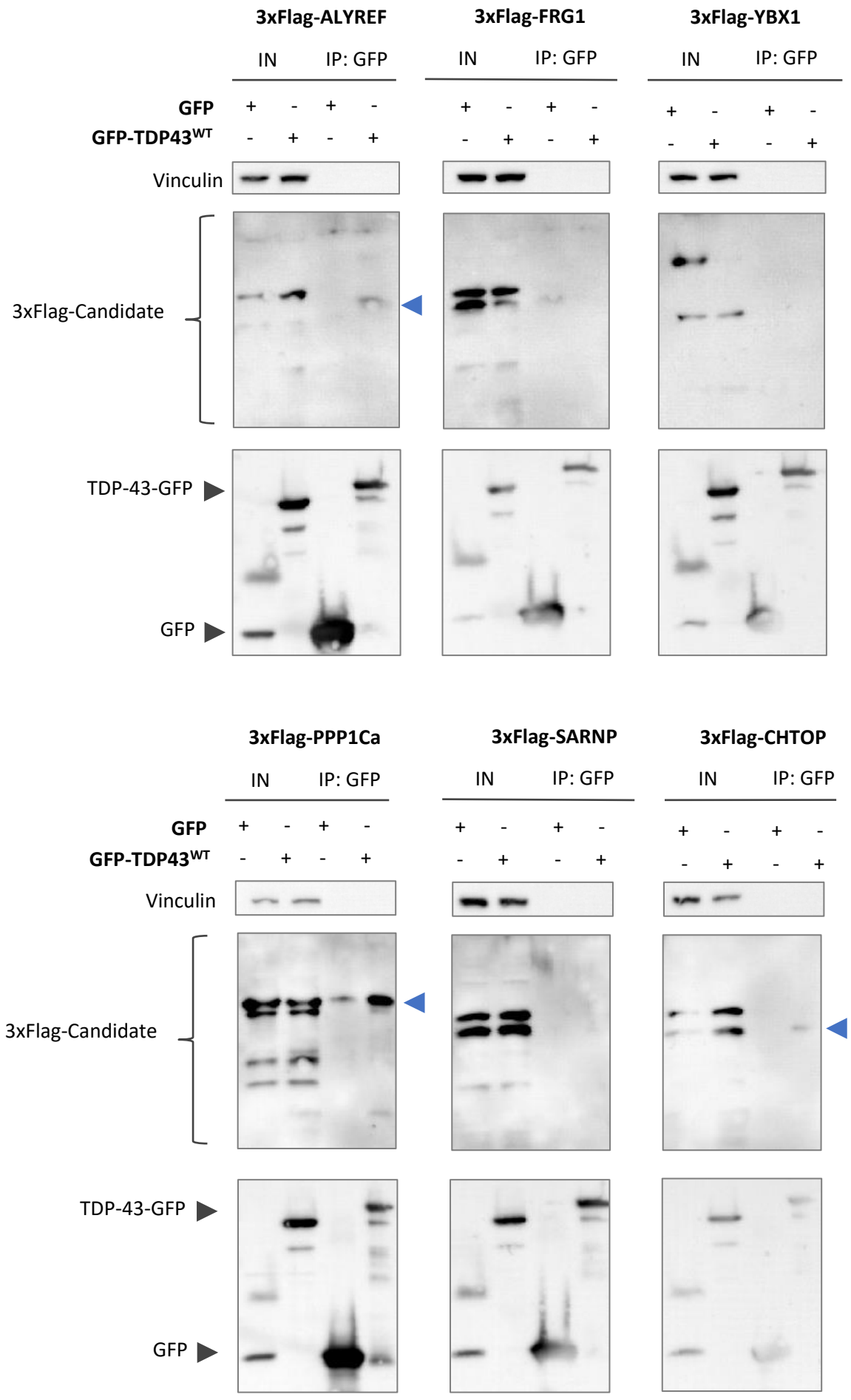


Figure S2 (caption on next page)

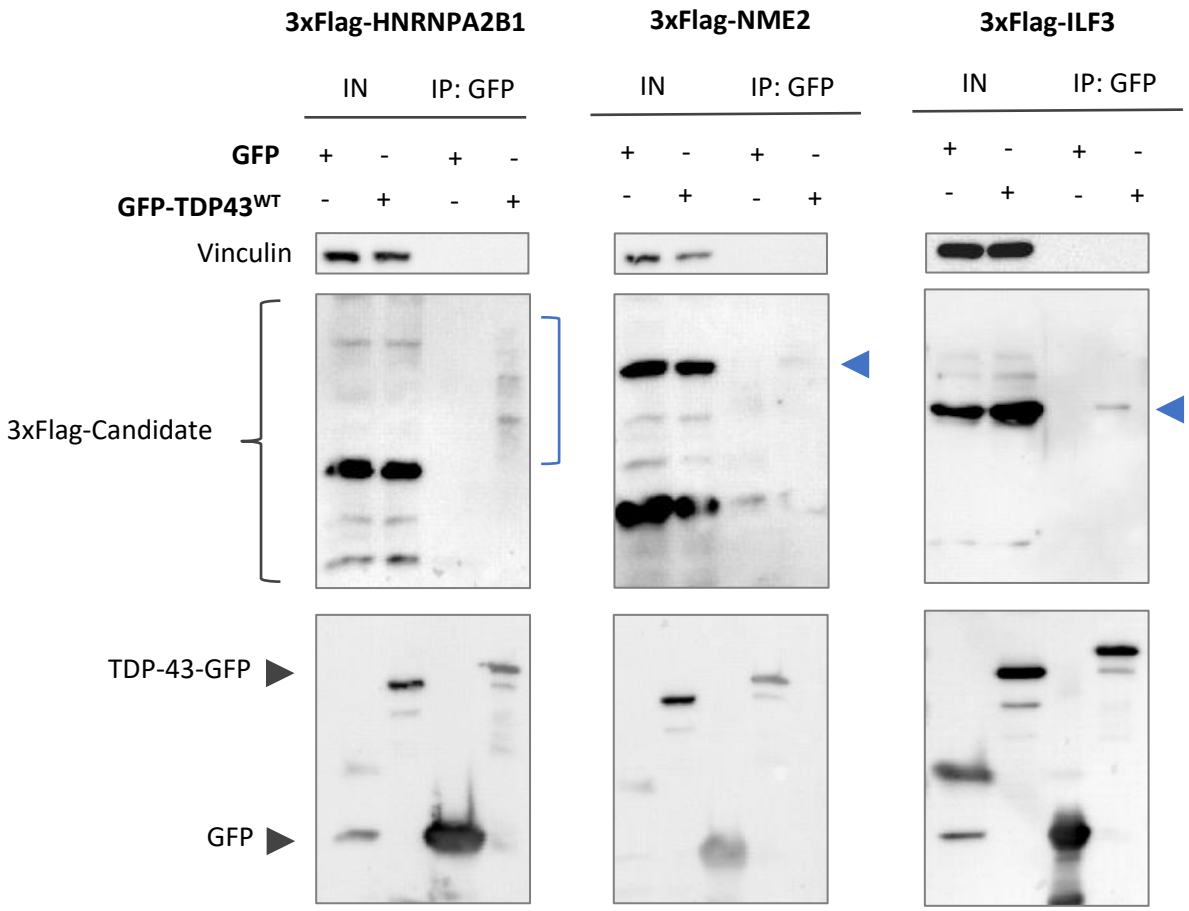


Figure S2: Validation of hit interaction with TDP-43^{WT}-GFP. Representative WBs showing co-overexpression GFP-IP experiments where GFP-TDP43^{WT} constructs were co-transfected with 3xFlag-hits. 3xFlag-tagged- ALYREF, FRG1, YBX1, PPP1Ca, SARNP, CHTOP, HNRNPA2B1, NME2, and ILF3 were co-transfected with TDP-43^{WT}-GFP expressing construct in HEK293T cells, and incubated for 48 hours to allow for expression. Blue arrows indicate expression of 3xFlag-hit. Blue curved line indicates expression of 3xF-HNRNPA2B1 isoforms. Top WB : probed using vinculin antibody. Middle WB: probed using Flag antibody. Bottom WB: Probed using GFP antibody.

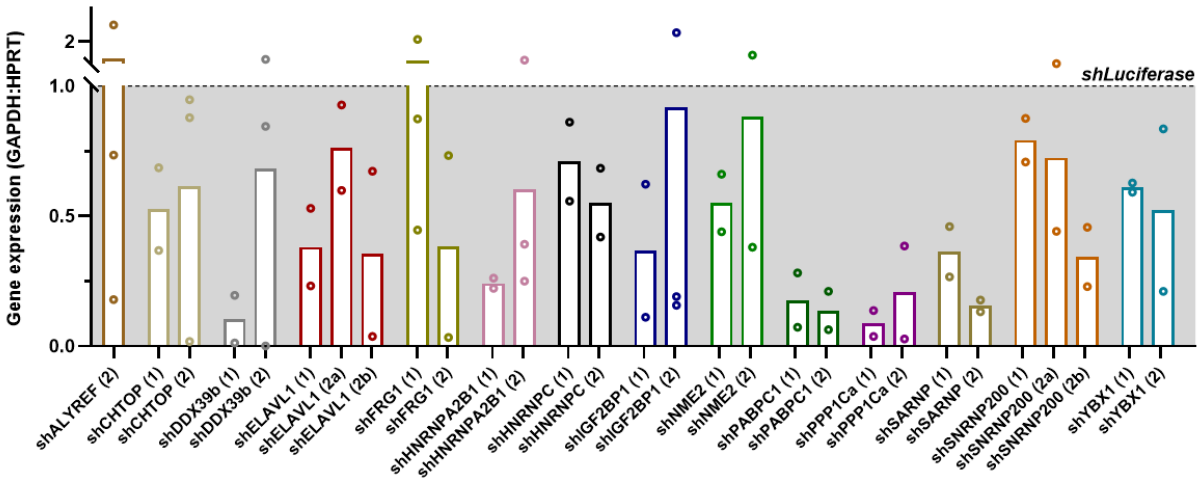
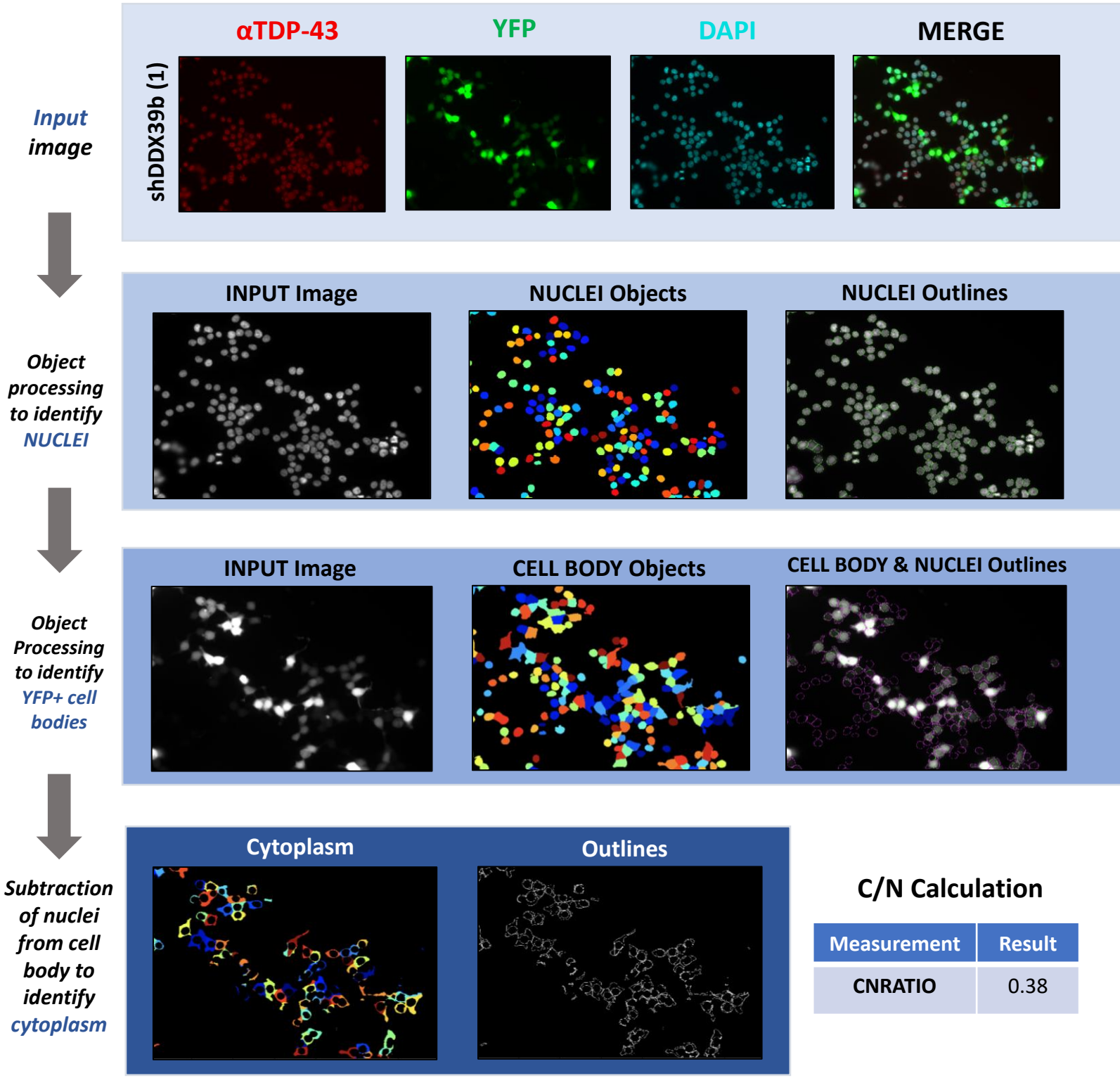


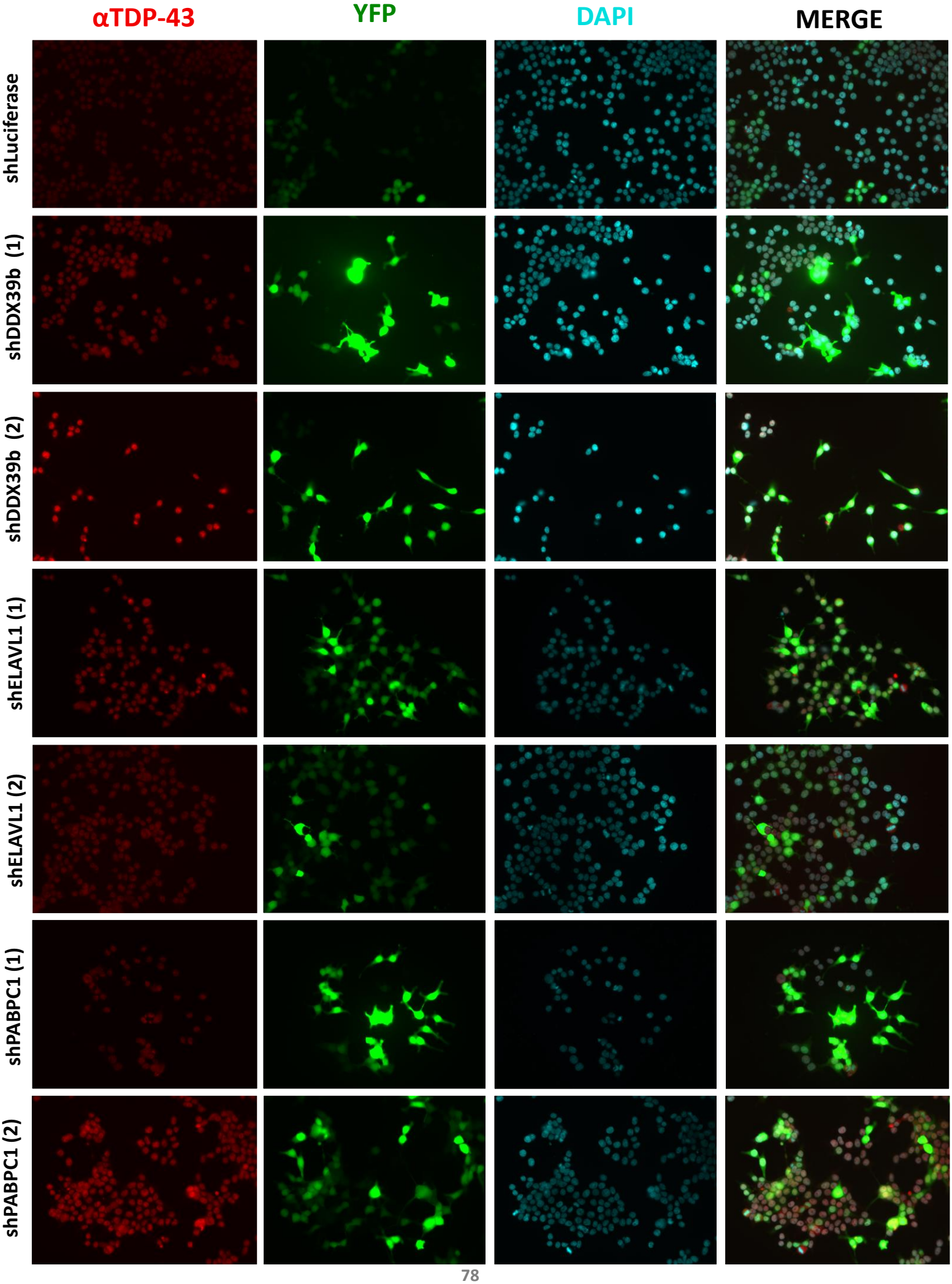
Figure S3: Knockdown efficiency of shRNAs against top hits as assessed by qPCR. (n = 2)* Experiments have been plotted as individual data points. Bars represent mean of experiments. Gene expression ((2^{-ddCT})) of shRNA target is normalized to average of housekeeping genes (GAPDH and HPRT). Expression of non-targeting shRNA (shLuciferase) is shown as dotted line. Colours of bars are unique to each gene target.
*qPCR experiments were repeated a minimum of two times, unless otherwise indicated by data points.



C/N Calculation

Measurement	Result
CNRATIO	0.38

Figure S4: Cell Profiler pipeline for analyzing microscopy images. [Top panels] Input images of shDDX39b-transfected HEK293T cells. Cells have been stained using TDP-43 antibody and imaged using 647 nm, EGFP and DAPI channels. Input image channels have been split and false coloured as the following: αTDP43 (red), YFP (green), DAPI (cyan). Merged image of all channels is shown.
[Row 2-4] Cell profiler program image analysis and object processing. Visual representation of Cell Profiler program pipeline used to identify cells, nuclei and cytoplasm from inputted images. Each colour represents an individual object (nuclei, cell body or cytoplasm). Average cytosol-to-nuclear intensity ratio of αTDP-43 was calculated and displayed.


αTDP-43
YFP
DAPI
MERGE
shLuciferase
shDDX39b (1)
shDDX39b (2)
shELAVL1 (1)
shELAVL1 (2)
shPABPC1 (1)
shPABPC1 (2)

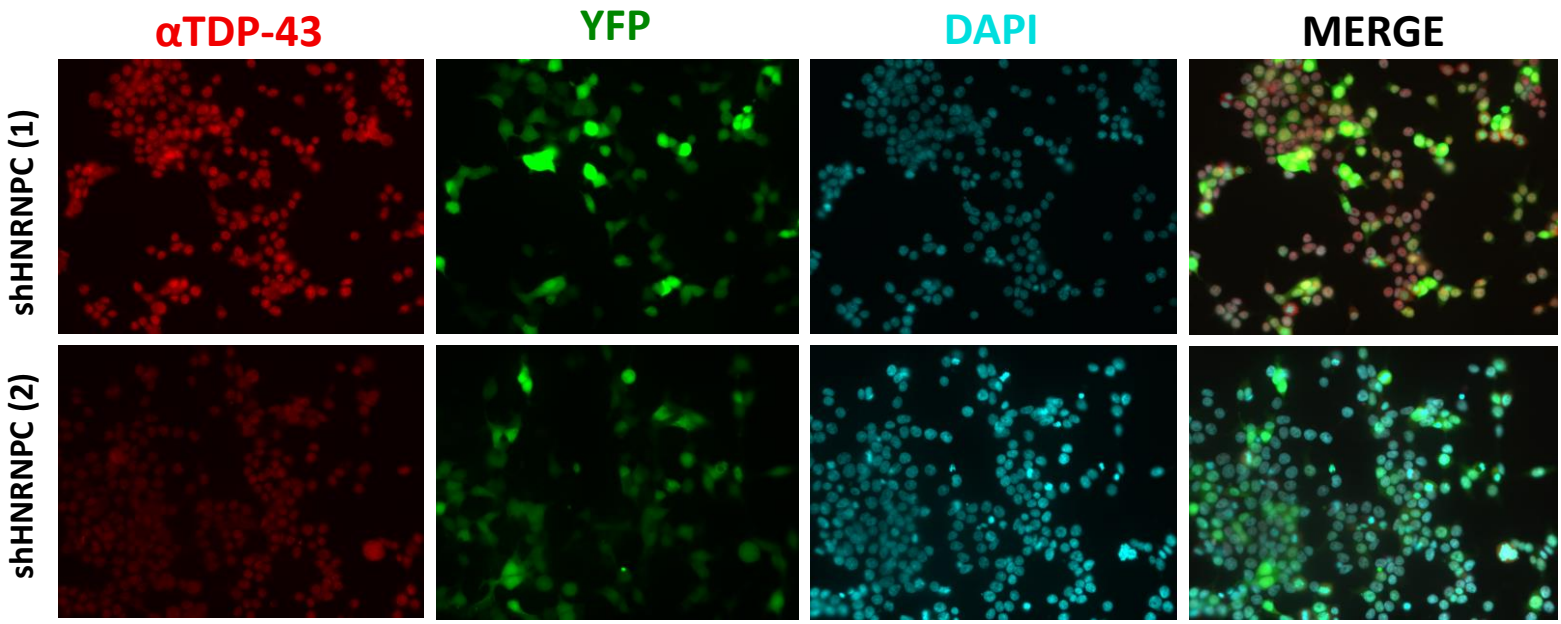


Figure S5: Assessing TDP-43 localization upon shRNA-hit knockdown via microscopy. Representative immunofluorescent microscopy images. HEK293T cells were transfected with shLuciferase, shDDX39b, shELAVL1, shHNRNPC and shPABPC1, with expression for 72 hours. DAPI (blue), TDP43 (red), EGFP (green), and merge of images is shown.

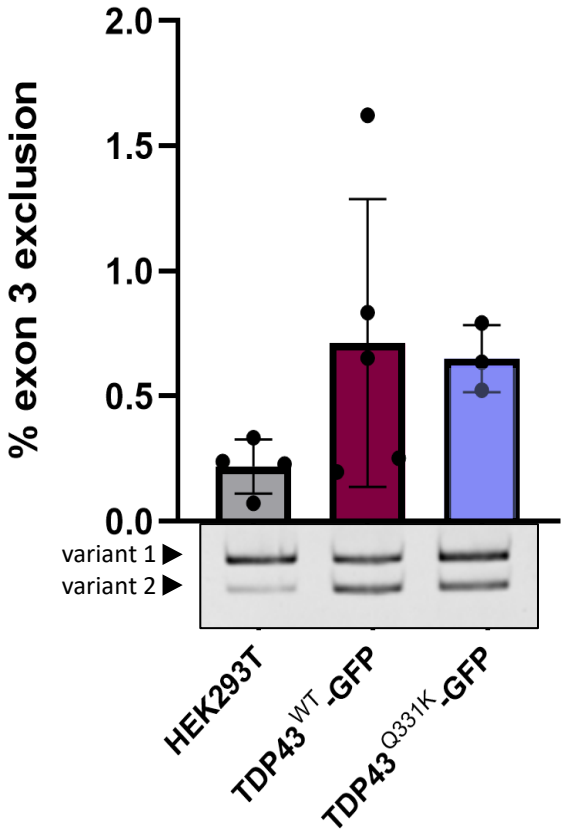


Figure S6: RT-PCR based splicing assay shows no significant TDP-43 loss of function between GFP-KI cells. HEK293T (n = 4), TDP-43^{WT}-GFP (n = 5) and TDP-43^{Q331K}-GFP (n = 3) cells were assayed in *POLDIP3* exon 3 splicing assay which shows trending increase in exon 3 exclusion, suggestive of TDP-43 loss of function with GFP-KI cells, but no significant differences between cell types. Graph is showing % of *POLDIP3* exon 3 exclusion, quantified from TAE PAGE gel bands (shown below graph). Bars represent mean of experiments. Error bars indicate standard deviation. Data was analyzed using one-way ANOVA with post-hoc analysis for multiple comparisons.

13. TABLES

TABLE 1: PRIMERS FOR TDP-43 GFP KI CELL CHARACTERIZATION

PRIMER NAME	SEQUENCE
QPCR TARDBP - F	CTGCTTCGGTGTCCCTGT
QPCR TARDBP - R	ATGGGCTCATCGTTCTCATC
QPCR SFGFP - F	ATGACGGGACCTACAAGACG
QPCR SFGFP - R	TGTGTCCGAGAATGTTTCCA
QPCR ADD2 - F	GGGGAGTGGAGAACCTCATC
QPCR ADD2 - R	CATAGGCCCAAAGGTGCTC
QPCR ATXN2 - F	TGAATGGCGAACACAAAGAG
QPCR ATXN2 - R	CAAGTTCCTCATTGGCTGTG
QPCR CDK6 - F	AAAATCTTGGACGTGATTGGA
QPCR CDK6 - R	TGGGAAGGGCAACATCTCTA
QPCR BBC3 - F	AAGAGCAAATGAGCCAAACG
QPCR BBC3 - R	CAAACGAGCCCCACTCTC
QPCR BAX - F	CAAGACCAGGGTGGTTGG
QPCR BAX - R	CACTCCCGCCACAAAGAT
QPCR CDKN1A - F	CTGGTACCCTCCTGGCTCTT
QPCR CDKN1A - R	CCCAGTGCAGGTCAGAGG
QPCR GAPDH – F	CGACCACTTTGTCAAGCTCA
QPCR GAPDH – R	TTACTCCTTGGAGGCCATGT
QPCR HPRT1 – F	GACCAGTCAACAGGGGACAT
QPCR HPRT1 – R	GTGTCAATTATATCTTCCACAATCAAG
QPCR STMN2 – F	CAATGGCTAAAACAGCAATGG
QPCR STMN2 – R	AGCAAGAGCAGATCAGTGACAG

TABLE 2: MASS SPECTROMETRY HITS

TDP-43^{WT} INTERACTORS (GENE NAME)	
ALYREF*	NHP2
BUB3	NME2
DDX39B*	NOLC1
DLD	PABPC1*
EIF4A3	PDIA6
ELAV1*	PPP1CA
FRG1*	PPP1CA
HNRNPA0	PPP1CC
HNRNPC*	PTBP1
HNRNPDL	RPL36A
HNRNPM	RPS24
IGF2BP1	RPS5
IGF2BP2	SAP18
IGF2BP3	SARNP*
ILF3*	SNRNP200*
KIF5B	TKT
MOV10	YBX1*

TDP-43^{WT} INTERACTORS (GENE NAME)	
CHTOP*	NOP56
COIL	NOP58
DKC1	PABPN1
ERH	SNU13
HNRNPA2B1*	TOE1
ILF2	VIM

***Indicates member of validation shortlist.*

TABLE 3: TOP HITS: PROTEIN NAME AND MOLECULAR WEIGHT

PROTEIN NAME	MOLECULAR WEIGHT (KDA)
ALYREF	25-35
CHTOP	25
DDX39B	50
ELAVL1	37
FRG1	25-37
HNRNPA2B1	36-38
HNRNPC	35-43
ILF3	45 (MONOMER) 90-110 (HETERODIMER)
NME2	41
PABPC1	70
PPP1CA	38
SARNP	29
SNRNP200	200
YBX1	50

TABLE 4: LOCALIZATION OF 3X-FLAG TAG

PROTEIN NAME	3X-FLAG LOCALIZATION (C/N TERMINUS)
ALYREF	C
CHTOP	C
DDX39B	N
ELAVL1	N
FRG1	C
HNRNPA2B1	N
HNRNPC	N
ILF3	N
NME2	N
PABPC1	N
PPP1CA	N
SARNP	C
YBX1	C

TABLE 5: hORFEOME LIBRARY ENTRY VECTORS FOR 3X-FLAG-HIT CONSTRUCT CLONING

Gene Name	Presence of stop codon	Antibiotic	Plate ID	Position Number	Additional cloning required?	Vector backbone	Sequencing primer	Approx. size (bp)
ALYREF	No	Kan	OCAA99	C3	No	pENTR221	T7 promoter	3317 bp
CHTOP	No	Spect	81046	F7	No	pDONR223	T7 promoter	5752 bp
DDX39b	Yes	Kan	OCAA22	D2	No	pENTR221	(-21)M13F	3831 bp
ELAVL1	Yes	Kan	OCAA55	G4	No	pENTR221	(-21)M13F	3524 bp
FRG1	No	Spect	OCAB45	B4	No	pDONR223	T7 promoter	3320 bp
HNRNPA2B1	No	Spect	81040	H6	Yes – add stop codon	pDONR223	(-21)M13F	6064 bp
HNRNPC	No	Spect	81060	G6	Yes – add stop codon	pDONR223	(-21)M13F	5923 bp
IGF2BP1	Yes	Spect	OCAB28	H4	No	pENTR223.1	(-21)M13F	6042 bp
ILF3	No	Spect	81014	C3	Yes – add stop codon	pDONR223	(-21)M13F	7687 bp
PABPC1	Yes	Kan	OCAA42	B2	No	pENTR221	(-21)M13F	4454 bp
PPP1CA	No	Spect	81062	F7	Yes – add stop codon	pDONR223	(-21)M13F	5995 bp
SARNP	No	Spect	81030	E8	No	pENTR221	T7 promoter	3176 bp
YBX1	No	Kan	OCAA74	A5	No	pENTR221	T7 promoter	3518 bp
NME2	No	Spect	RRID:Addgene_23392		Yes – add stop codon	pDONR223	T7 promoter	5005 bp

TABLE 6: 3X-FLAG-HIT CLONING SITE DIRECTED MUTAGENESIS (SDM) PRIMERS

NAME	SEQUENCE (5' → 3')
NME2-Stop-Fwd	gtgctcatgactgggtctatgaatagtgccaactttcttgtaca
NME2-Stop-Rev	tgtacaagaaagttggcactattcatagacccagtcatgagcac
HNRNPA2B1-Stop-Fwd	ctaactacggtccaaTAgcccaactttcttgtacaaagttgg
HNRNPA2B1-Stop-Rev	ccaactttgtacaagaaagttgggcTAttggaccgtagttag
PPP1CA-Stop-Fwd	actttgtacaagaaagttgggCtatttcttggctttggcggaattgcggg
ILF3-Stop-Fwd	ctttttcacagactgctacggctatcatgattttgggtcttcctaGcccaactttcttgtacaaagt
HNRNPC-Stop-Fwd	cgccaatggcgaggatgactcttagcccaactttcttgtac

TABLE 7: OLIGOMERS SEQUENCES FOR SHRNA GENERATION

shRNA NAME	OLIGOMER
shALYREF-1	TGCTGTTGACAGTGAGCGCGGACGACATCATTA AACTGAATAGTGAAGCCACAGATGTATTCAGTT TAATGATGTCGTCCATGCCTACTGCCTCGGA
shALYREF-2	TGCTGTTGACAGTGAGCGCAGGATTTAAAAACT CATGTAATAGTGAAGCCACAGATGTATTACATG AGTTTTTAAATCCTATGCCTACTGCCTCGGA
shDDX39b-1	TGCTGTTGACAGTGAGCGCCAGCAGTTTAAAGA TTTTCAATAGTGAAGCCACAGATGTATTGAAAA TCTTTAAACTGCTGATGCCTACTGCCTCGGA
shDDX39b-2	TGCTGTTGACAGTGAGCGCCAGAATTTTTTTTTA ACAAAATAGTGAAGCCACAGATGTATTTTGTTA AAAAAAAATTCTGATGCCTACTGCCTCGGA
shELAVL1-1	TGCTGTTGACAGTGAGCGCTTGTACGGAATAGA TAATTAATAGTGAAGCCACAGATGTATTAATTA TCTATTCCGTACAAATGCCTACTGCCTCGGA
shELAVL1-2	TGCTGTTGACAGTGAGCGAGAGAACGAATTTGA TCGTCAATAGTGAAGCCACAGATGTATTGACGA TCAAATTCGTTCTCCTGCCTACTGCCTCGGA
shSARNP-1	TGCTGTTGACAGTGAGCGATGAAGATGATGAGA AACTGAATAGTGAAGCCACAGATGTATTCAGTT TCTCATCATCTTCAGTGCCTACTGCCTCGGA
shSARNP-2	TGCTGTTGACAGTGAGCGCGAAGATGATGAGAA ACTGAAATAGTGAAGCCACAGATGTATTTCAGT TTCTCATCATCTTCATGCCTACTGCCTCGGA
shPABPC1-1	TGCTGTTGACAGTGAGCGCCGGAACACACATTT ATTATTATAGTGAAGCCACAGATGTATAATAAT AAATGTGTGTTCCGATGCCTACTGCCTCGGA
shPABPC1-2	TGCTGTTGACAGTGAGCGCTGGTAACATCCTTT CATGTAATAGTGAAGCCACAGATGTATTACATG AAAGGATGTTACCAATGCCTACTGCCTCGGA
shFRG1-1	TGCTGTTGACAGTGAGCGATGTGTTTTTTTCTGA ATAAAATAGTGAAGCCACAGATGTATTTTATTC AGAAAAAAACACAGTGCCTACTGCCTCGGA
shFRG1-2	TGCTGTTGACAGTGAGCGCTCTGTGTTTTTTTCT GAATAATAGTGAAGCCACAGATGTATTATTCAG AAAAAAACACAGAATGCCTACTGCCTCGGA
shIGF2BP1-1	TGCTGTTGACAGTGAGCGATGGTGCTGTTGAGA TATTTTATAGTGAAGCCACAGATGTATAAAATA TCTCAACAGCACCACTGCCTACTGCCTCGGA
shIGF2BP1-2	TGCTGTTGACAGTGAGCGAGTCGGCCATTTAGT AATTTAATAGTGAAGCCACAGATGTATTAAATT ACTAAATGGCCGACGTGCCTACTGCCTCGGA
shYBX1-1	TGCTGTTGACAGTGAGCGCGAAGAAGATAAAG AAAATCAATAGTGAAGCCACAGATGTATTGATT TTCTTTATCTTCTTCATGCCTACTGCCTCGGA

shYBX1-2	TGCTGTTGACAGTGAGCGAACCAAGGAAGATGT ATTTGTATAGTGAAGCCACAGATGTATACAAAT ACATCTTCCTTGGTGTGCCTACTGCCTCGGA
shSNRNP200-1	TGCTGTTGACAGTGAGCGCAGTAGTGAAGAAGA TTGAGAATAGTGAAGCCACAGATGTATTCTCAA TCTTCTTCACTACTTTGCCTACTGCCTCGGA
shSNRNP200-2	TGCTGTTGACAGTGAGCGCCAACACCTTTGATT TCATTAATAGTGAAGCCACAGATGTATTAATGA AATCAAAGGTGTTGATGCCTACTGCCTCGGA
shHNRNPC-1	TGCTGTTGACAGTGAGCGCTCCCATGTTCATTA ATTCATATAGTGAAGCCACAGATGTATATGAAT TAATGAACATGGGAATGCCTACTGCCTCGGA
shHNRNPC-2	TGCTGTTGACAGTGAGCGATAGTAGTTTTGTTA AGTCTTATAGTGAAGCCACAGATGTATAAGACT TAACAAAACTACTAGTGCCTACTGCCTCGGA
shPPP1CA-1	TGCTGTTGACAGTGAGCGCCGCGGCCATAGTGG ACGAAAATAGTGAAGCCACAGATGTATTTTCGT CCACTATGGCCGCGATGCCTACTGCCTCGGA
shPPP1CA-2	TGCTGTTGACAGTGAGCGCCGGCTACGAGTTCT TTGCCAATAGTGAAGCCACAGATGTATTGGCAA AGAACTCGTAGCCGTTGCCTACTGCCTCGGA
shNME2-1	TGCTGTTGACAGTGAGCGAACAGGCCTAATCTC TATGACATAGTGAAGCCACAGATGTATGTCATA GAGATTAGGCCTGTCTGCCTACTGCCTCGGA
shNME2-2	TGCTGTTGACAGTGAGCGACAGGGACACTGCAA GTAGGAATAGTGAAGCCACAGATGTATTCCTAC TTGCAGTGTCCCTGGTGCCTACTGCCTCGGA
shCHTOP-1	TGCTGTTGACAGTGAGCGACACCACCAAGATGT CTCTAAATAGTGAAGCCACAGATGTATTTAGAG ACATCTTGGTGGTGCTGCCTACTGCCTCGGA
shCHTOP-2	TGCTGTTGACAGTGAGCGCCCACATTGATAATT TAGTAAATAGTGAAGCCACAGATGTATTTACTA AATTATCAATGTGGATGCCTACTGCCTCGGA
shHNRNPA2B1-1	TGCTGTTGACAGTGAGCGAACAGATTACAATTC TATTTTATAGTGAAGCCACAGATGTATAAAATA GAATTGTAATCTGTCTGCCTACTGCCTCGGA
shHNRNPA2B1-2	TGCTGTTGACAGTGAGCGCACAATTCTATTTTA AATATAATAGTGAAGCCACAGATGTATTATATT TAAAATAGAATTGTATGCCTACTGCCTCGGA

TABLE 8: QPCR PRIMER NAME AND SEQUENCE

QPCR PRIMER NAME	PRIMER SEQUENCE
QPCR-ALYREF-FWD	CAGCAGACGTGCACTTTGAG
QPCR-ALYREF-REV	TGACGTGACAAGCTGAATGTT
QPCR-CHTOP-FWD	GAAGAGGTGGTGTTCGAGGT
QPCR-CHTOP-REV	TTCCCCGACCTATCATACCC
QPCR-DDX39B-FWD	GAAGAGGTGCTGAAGAAGAACTGCC
QPCR-DDX39B-REV	CACTGAACATCATGACCTGCTTCTCG
QPCR-ELAVL1-FWD	CAGGCGCAGAGATTCAGG
QPCR-ELAVL1-REV	GGTTGTAGATGAAAATGCACCA
QPCR-SARNP-FWD	CAGCTGGAACTGGAACCAC
QPCR-SARNP-REV	TTCATCAGGCAATCCCAAAG
QPCR-PABPC1-FWD	CAGGGTGCCAGACCTCAT
QPCR-PABPC1-REV	AGCTGGTCTCATAGTACTAAATGGTG
QPCR-FRG1-FWD	CAGGGGACATAGAAGCAAAAA
QPCR-FRG1-REV	TCTCTTTCAGCACAGGATCTAATCT
QPCR-IGF2BP1-FWD	ACAAAACAGACCCAGTCCAAG
QPCR-IGF2BP1-REV	GCACACTGATGGCTTTTTCA
QPCR-YBX1-FWD	GGAGGGTGCTGACAACCA
QPCR-YBX1-REV	GCTGTCTTTGGCGAGGAG
QPCR-SNRNP200-FWD	GCTCCTTGGTGCAGGAGAT
QPCR-SNRNP200-REV	CAGCAACAGTGATGCCATAAG
QPCR-HNRNPC-FWD	TGACTTTCAACGGGACTATTATGA
QPCR-HNRNPC-REV	TCCTGATACACGCTGACGTT
QPCR-PPP1CA-REV	GACAGCGAGAAGCTCAACCT
QPCR-PPP1CA-FWD	CGCGGATCTCGTTCTCTG
QPCR-HNRNPA2B1-REV	CGCGGAGATCTCTCTCATCT
QPCR-HNRNPA2B1-FWD	GAACTGTTCCTTTTCTCTCTCCA
QPCR-NME2-FWD	GAAAACTCGACCGCACTTTAG
QPCR-NME2-REV	GATGATCTCGCCCACCAG

TABLE 9: SPLICING ASSAY PRIMERS

PRIMER NAME	PRIMER SEQUENCE
Fw-POLDIP3-e3	GCTTAATGCCAGACCGGGAGTTG
Rv-POLDIP3-e3	TCATCTTCATCCAGGTCATATAAATT

TABLE 10: ANTIBODIES

ANTIBODY	COMPANY	CATALOG NO.	SPECIES	DILUTION
ALEXA FLUOR 488, SECONDARY	Thermo Fisher Scientific	A32814	αGoat	IF: 1:1,000
ALEXA FLUOR 488, SECONDARY	Thermo Fisher Scientific	A11001	αMouse	IF: 1:1,000
ALEXA FLUOR 647, SECONDARY	Thermo Fisher Scientific	A21244	αRabbit	IF: 1:1,000
DDDDK (FLAG TAG)	Abcam	AB1257	Goat	IF: 1:1,000
FLAG	University of Ottawa	Biobar	Rat	WB: 1:1,000
FLAG M2	Sigma-Aldrich	F1804	Mouse	WB: 1:2,000
GFP	Abcam	AB6556	Rabbit	WB: 1:2,000, 1:5,000
GFP	Santa Cruz Biotechnology	SC-9996	Mouse	WB: 1:2,000
SNRNP200	Sigma-Aldrich	HPA029321	Rabbit	WB: 1:5,000
TARDBP (TDP-43)	Cedarlane/Proteintech	10782-2-AP	Rabbit	IF: 1:1,000 WB: 1:2,000, 1:5,000
VINCULIN	Sigma-Aldrich	V9131	Mouse	WB: 1:20,000
RAT-HRP	Thermo Fisher Scientific	31470	αRat	WB: 1:10,000
MOUSE-HRP	Cedarlane/Jackson ImmunoResearch	715-035-150	αMouse	WB: 1:10,000
RABBIT-HRP	Cedarlane/Jackson ImmunoResearch	711-035-152	αRabbit	WB: 1:10,000