UNIVERSITY OF CALIFORNIA SAN DIEGO

Non coding RNA MALAT1 regulates TDP-43 binding activity in neurodegenerative disease conditions

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Biochemistry and Molecular Biophysics

by

Adarsh Balaji

Committee in charge:

Professor Colleen A. McHugh, Chair
Professor Alexis Komor
Professor Miles Wilkinson
Professor Brian Zid

2024

The Dissertation of Adarsh Balaji is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ALS	Amyotrophic Lateral Sclerosis
AS	Alternative splicing
BARD1	BRCA1 Associated RING Domain 1
CLIP	Cross-linked Immunoprecipitation
CLIP-seq	Cross-linked immunoprecipitation coupled with sequencing
CSNK1E	Casein Kinase 1 E
CSTF2	Cleavage Stimulation Factor Subunit 2
DAPI	4’,6-diamidino-2-phenylindole
DNA	Deoxyribonucleic Acid
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
HMGB2	High Mobility Group Box 2
hnRNP	Heterogeneous Nuclear Ribonucleoprotein
in4	Intron-4
in16	Intron-16
lncRNA	Long Non-Coding RNA
MADD	MAPK activating death domain protein
MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1
NEAT1	Nuclear-Enriched Abundant Transcript 1
MPP^+	1-methyl-4-phenylpyridinium
mRNA	Messenger RNA
miRNA	micro-RNA
PD	Parkinson’s Disease

PPFIA3	PTPRF Interacting protein alpha 3
RAP-MS	RNA Antisense Purification coupled with Mass Spectrometry
RNA	Ribonucleic Acid
RNA-seq	Ribonucleic Acid sequencing
SAT1	Spermine/Spermidine Acetyltransferase 1
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis
SLC1A5	Solute Carrier Family 1 Member 5
TARDBP	TDP-43 Protein Encoding Gene
TDP-43	Transactive Response DNA Binding Protein 43
UG	Uracil-Guanine
UTR	Untranslated Region

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my thesis advisor, Colleen McHugh, for the support and mentorship that I have received in my degree. Colleen was willing to take me on as a graduate student even though I had no prior experience working with RNA or molecular biology, and she has strongly supported me in my research endeavors. She has helped me in writing my grants, publications, and given me helpful advice and mentorship for my career goals as well. I would not be the scientist that I am today without her guidance, and I really appreciate her mentorship over these last five years.

I also want to acknowledge my postdoctoral mentor during my undergraduate research career, Dr. Stefanie Volland. Stefanie was my first scientific mentor when I joined the Williams Lab, and she helped my initial development as a scientist. Through Stefanie, I learned how to critically think about hypotheses, design and validate experiments with the appropriate controls, and how to research the existing literature to come up with knowledge gaps that can potentially be answered by our experimental findings. Stefanie helped instill within me a belief and interest in pursuing a research career and a scientific degree at a higher education level, and I would not have had the confidence or skill-set to apply for a PhD program without her support.

I would also like to thank my committee members – Dr. Alexis Komor, Dr. Itay Budin, Dr. Partho Ghosh, Dr. Miles Wilkinson, Dr. Brian Zid - for their guidance and oversight during my degree. Dr. Itay Budin and Dr. Partho Ghosh were initial members of my committee and were present during my second and third year examinations. Unfortunately, they were unable to participate in my final examinations, but I really appreciate the help and support that they provided during my initial exams. Dr. Brian Zid was a later addition to my final examination committee and I appreciate his willingness to make time and effort to provide guidance and

suggestions for my thesis work. I also thank Dr. Alexis Komor and Dr. Miles Wilkinson for being on my committee observing my work from the initial stages to the final examination.

My experiments would not have been successful without the collaborative efforts in my research lab. I really appreciate the guidance and support from my fellow graduate students in the McHugh Lab – Aileen, Tong, Nhu, Simone, Lauren, Jonathan. Aileen and Tong have been mentors for me since I joined, and their willingness to listen to my various hypotheses and (quite frankly) crazy ideas at all times was a blessing during my degree. I also want to thank my undergraduate student, Raul Johnson, for helping me out with my experiments and contributing to my research. I would also like to thank my collaborators, Ali and Zeinab, for their contributions towards my experiments and helping us with fluorescence microscopy visualization.

I came to San Diego to pursue my PhD and ventured into a new city without knowing anyone. Thankfully, I was able to meet some amazing scientists (and significant others) who helped ease this transition, and have become close and dear friends to me. Whether we were discussing science, playing board games, or lampooning as adventurers in a classic DnD game, we have provided support and encouraged each other from the beginning and through the pandemic. I would like to thank Emily, Andrew, Alysia, Freddy, Lannah – and a special shout out to the 'Kiosco Crew', Chee, Sean, and Sally - for the camaraderie and experiences we have shared together these last few years.

I also want to thank my best friends from college, who have been here for me since day 1. They have inspired me to continue to work harder, to be more successful, and they have always been here for me when I've been at my lows and at my highs. I want to specifically thank Denver, Ashwin, Sohom and Anjali for the memories and support that they have given me over

the last 10 years, and I look forward to seeing how our friendships grow stronger over the coming years.

Finally, I would like to acknowledge the massive encouragement and love that I have received from my family - Amma, Appa, Abhi and Sravya. To my parents, thank you for believing in me and for the countless sacrifices you made to ensure I had the opportunity to pursue my dreams. You both decided to move to the US to support my brother and I in college, and to ensure we could transition from being international students to US citizens. Your love and encouragement have been my constant source of strength, especially during the most challenging times. My brother (and his better half, Sravya) has been a constant role model for me in my life – if I am half as successful as he is, I consider that to be a resounding success. I hope to make you all proud with my future endeavors in the scientific community, and I dedicate this thesis to your endless support, understanding, and love.

Chapter 2, in full, has been submitted for publication of the material as it may appear in the Journal of Biological Chemistry, and is currently in revision. This material includes author contributions from Aileen C. Button, Jonathan Zhu, Lauren Ellis, Ellen Lavorando, Ethan L. Ashley, Raul Johnson, Einollah Sarikhani, Zeinab Jahed, and Colleen A. McHugh. The dissertation author was the primary researcher and author of this material.

Chapter 3, in full, is currently in preparation for submission for publication of the material. This material includes author contributions from Raul Johnson, Jonathan Zhu, Lauren Ellis, Simone Hall, and Colleen A. McHugh. The dissertation author was the primary researcher and author of this material.

VITA

2019 Bachelor of Science in Biochemistry, Minor in Biomedical Research, University of California Los Angeles

2021 Master of Science in Chemistry, University of California San Diego

2024 Doctor of Philosophy in Biochemistry and Molecular Biophysics, University of California San Diego

PUBLICATIONS

Jiang, M., Paniagua, A.E., Volland, S., Wang, H., **Balaji, A.**, Li, D.G., Lopes, V.S., Burgess, B.L. and Williams, D.S., 2020. “Microtubule motor transport in the delivery of melanosomes to the actin-rich apical domain of the retinal pigment epithelium”. Journal of Cell Science, 133(15).

Balaji, A., Button, A.C., Zhu, J., Ellis, L., Lavorando, E., Ashley, E.L., Johnson, R., Sarikhani, E., Jahed, Z., McHugh, CA. “MALAT1 RNA levels affect cell viability and modulate TDP-43 binding to mRNA in the Nucleus”. Journal of Biological Chemistry. (In Review)

Balaji, A., Johnson, R., Zhu, J., Ellis, L., Hall, S., McHugh, C.A. "MALAT1 lncRNA aids efficient alternative splicing through sequence specific interactions to splicing factor proteins and pre-mRNA targets”. Manuscript in preparation for submission.

ABSTRACT OF THE DISSERTATION

Non coding RNA MALAT1 regulates TDP-43 binding activity in neurodegenerative disease conditions

by

Adarsh Balaji

Doctor of Philosophy in Biochemistry and Molecular Biophysics

University of California San Diego, 2024

Professor Colleen A. McHugh, Chair

Chapter 1 is an introduction to the work detailed in this thesis and details the background of the main biological molecules that are studied. The recent discovery of novel long-non coding RNAs (lncRNAs) reveals a new class of molecules with immense biological relevance and

promising candidates for therapeutic targeting. Classified as RNA molecules longer than 200 nucleotides in length, lncRNAs have been found to be dysregulated in various disease conditions, and their return to steady-state levels through transcript level manipulation has been observed to induce cell and tissue survival. Of great interest is the lncRNA MALAT1 which is known to be dysregulated in various cancer tissues and neurodegenerative diseases. Localized within the nucleus, MALAT1 is involved in a variety of important biological processes including mRNA transcriptional regulation, alternative splicing, and stabilizing nascent nuclear speckles. A detailed analysis of direct protein binders to MALAT1 reveals TDP-43 as a strong binder with common functionalities present. Both MALAT1 and TDP-43 are dysregulated in neurodegenerative diseases, and their common functionalities overlap in alternative splicing and transcript regulation. An increase in MALAT1/TDP-43 binding is also observed in neurodegenerative diseased tissues, but the reason for this interaction is yet to be explored fully in literature. Thus, it is important to study the common functions of MALAT1 and TDP-43 in the context of neurodegeneration and understand if the MALAT1/TDP-43 complex plays a role in regulating biological mechanisms for cell survivability, which can then be used for therapeutic advancements in this field.

Chapter 2 is focused on the regulation of TDP-43 binding to mRNA targets by MALAT1. The increase in MALAT1/TDP-43 interaction in neurodegenerative disease tissues indicates a potential function for this interaction. Given TDP-43's known function of regulating gene expression level, we hypothesized that MALAT1 may play a role in regulating TDP-43 function and gene regulation. We found that altering TDP-43 expression alters MALAT1 expression, and disruption of the steady-state level of either transcript induces cell death in neuroblastoma samples. TDP-43 binds to the 3'UTR of mRNA genes, and reduced MALAT1 expression can

increase the binding ability of TDP-43 to target 3'UTR genes in HEK293 and SH-SY5Y cells. This suggests that MALAT1 regulates the binding activity of TDP-43 to target genes in the nucleus. This hypothesis is extended in a Parkinson's Disease model through the induction of MPP^+ to SH-SY5Y cells, and a similar alteration in TDP-43 binding and gene expression regulation is observed in altered MALAT1 conditions. From the collected data, we can infer the existence of a coordinated network of RNA-protein binding that mediates gene expression levels and ensures homeostatic concentrations in cell. Dysregulation of the MALAT1 lncRNA results in aberrant TDP-43 function, which has severe consequences for cell viability.

Chapter 3 details the role of MALAT1 in regulating alternative splicing through sequence-specific binding interactions. Alternative splicing (AS) regulation is a common function of both TDP-43 and MALAT1, and the increase in TDP-43/MALAT1interaction may affect gene expression level and molecular pathway regulation through changes in AS. We hypothesized that MALAT1 and TDP-43 may concurrently regulate gene expression of a common splice target, the spermine/spermidine acetyltransferase 1 (SAT1) pre-mRNA. Further investigation revealed the existence of overlapping sequence-specific binding regions between MALAT1/SAT1 and MALAT1/TDP-43 that regulated SAT1 AS. These regions help efficiently recruit SAT1 pre-mRNA to TDP-43, with MALAT1 acting as a bridge to direct localization. This is a novel proposed mechanism for AS regulation by lncRNA. We then extended this hypothesis to another splicing factor and pre-mRNA target and observed a similar recruitment phenomenon for the MALAT1/CSTF2/PPFIA3 interactome. We hence present a model in which MALAT1 can recruit both pre-mRNA and splicing factor protein via an overlapping binding region with distinct sequence-specific targeting for efficient AS.

Chapter 1 : INTRODUCTION

1.1 Long non-coding RNA

The central dogma of molecular biology is a classical theory that describes the flow of genetic information in biological organisms from DNA to RNA, to proteins [1]. Often observed as a forward process, this created a controversial belief that the primary function of RNA was to be translated to proteins – that it would predominantly exist and function as messenger RNA. However, this belief was challenged by observations that clashed with a unidirectional flow of genetic information – for example, bacteria were known to undergo 'reverse transcription' to convert RNA to DNA [2]. Such exemptions from the classical theory thus posed a question about the fundamental nature of RNA molecules – could they function as more than messenger RNA? Did they have structure and intrinsic function without undergoing translation?

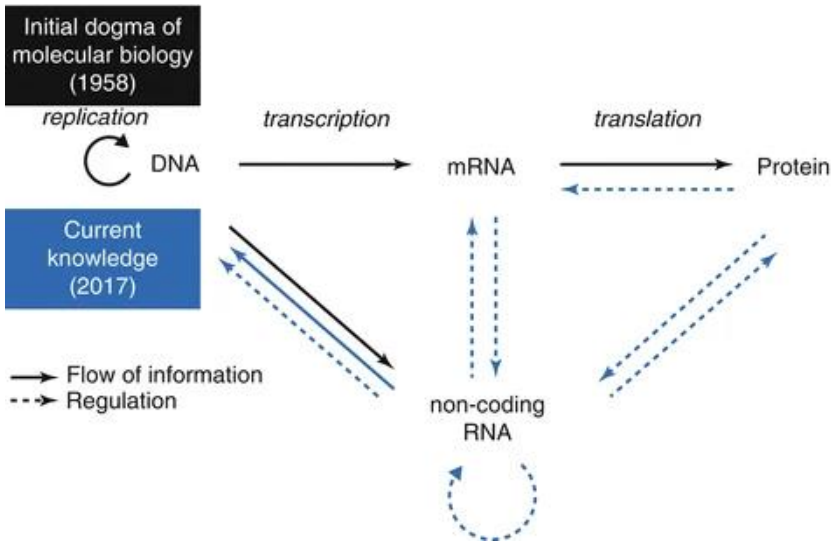


Figure 1.1 - The Central Dogma of Biology

This figure was originally published in History, Discovery, and Classification of lncRNAs [2].

To obtain a better understanding of the identification and functions of RNA within the human genome, large scale sequencing projects were required. The Encyclopedia of DNA Elements

(ENCODE) Project started in 2003 with the goal of defining elements of the human genome that had previously been discovered from the Human Genome Project but had yet to be annotated [3]. ENCODE sought to annotate the human genome in its entirety and would soon expand to include genomes of other organisms. Over a span of 15 years, the number of protein coding genes identified by the ENCODE project remained roughly the same, while thousands of novel regions (previously identified as "junk DNA") were now annotated as various forms of non-coding RNA – RNAs that RNAs that were only transcribed and were not translated into proteins.

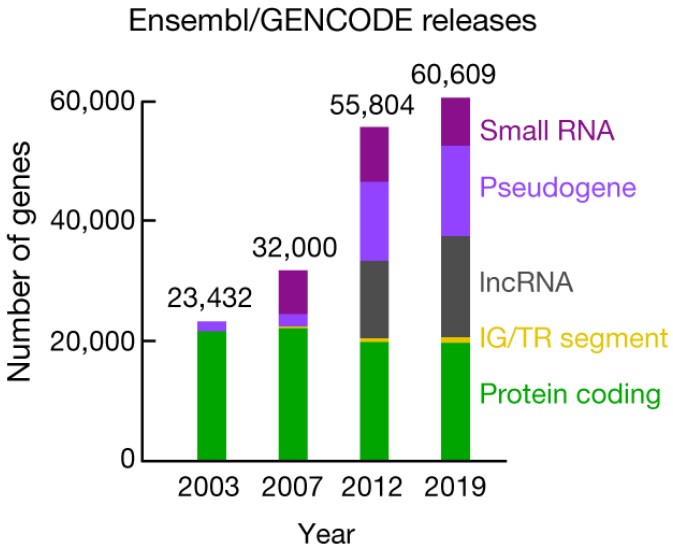


Figure 1.2 - Gene annotations reveals an increasing amount of non-coding RNA
This figure was adapted from Perspectives on ENCODE [3].

Examples of non-coding RNA have already been studied in literature. These included the common tRNAs, rRNAs, snRNAs, and others. These ncRNAs were typically small in length, and their function was often assumed to be interacting with protein subunits to form larger macro complex structures for biological function. The novel sequences annotated by the ENCODE project revealed many RNAs of more than 200nt in length, which were subsequently referred to as long noncoding RNAs (lncRNAs). More than 16,000 genes have been conservatively

annotated in the human genome to encode more than 28,000 diverse lncRNA transcripts [4], which continue to be researched for novel functional and structural analysis.

lncRNAs have a wide variety of functions, and their role in biological pathways are often from their interaction with other biomolecules (RNA, DNA, or proteins) to form larger macrocomplexes. The composition and structural organization of these large complexes allow lncRNAs to function in many biological pathways. lncRNAs are often involved in epigenetic, pre and post-transcriptional processing of mRNA, and translational regulation as well. This diverse functionality also indicates that lncRNAs can be localized in either the cytoplasm or the nucleus of cells. One common method through which lncRNAs are involved in metabolic pathways is through recruitment of other biological molecules by acting as a molecular scaffold or sponge. Many lncRNAs are known to be involved in large macrocomplex assembly through interaction with other RNA molecules or proteins. The colocalization of lncRNA NEAT1 and MALAT1 to nuclear paraspeckles helps recruit splicing factors to pre-mRNA targets and regulate efficient alternative splicing [5]. lncRNAs have also been studied in the context of microRNA (miRNA) sponging to influence physiological processes and regulate homeostatic gene expression [6].

Given their importance in maintaining biological pathways, the aberrant regulation and expression of lncRNAs is a hallmark of disease conditions and has been noted to positively correlate with disease progression. In addition to cancers, over 40% of human lncRNAs have been found to have spatiotemporal dynamic expression levels in different regions of brain tissue [7, 8]. This has been hypothesized to promote neuron differentiation and specialization of regional activity [9]. Thus, the functional study of these lncRNAs may prove helpful in

discovering new therapeutic modalities and targets through which disease progression can be mitigated.

1.2 MALAT1

The Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT1) lncRNA is one of the most studied lncRNAs in literature. Alternatively labelled as a nuclear-enriched abundant transcript 2 (NEAT_2), it is located on chromosome 11q13 and is downstream of another highly studied lncRNA NEAT1 [10]. The MALAT1 transcript is around 8.5Kb and results in the formation of a lncRNA with a cleaved 3' end, processed by RNAse P and Z. This 3'end structure, known as the MALAT1-associated small cytoplasmic RNA (mascRNA) folds similarly as tRNA to resemble a clover-like structure, and localizes to the cytoplasm. The processed MALAT1 transcript localizes to the nucleus.

MALAT1 was initially discovered to be strongly associated with metastasis in early-stage non-small cell lung cancer tissues and was annotated as an oncogene [11]. Further studies revealed high basal transcript level in many organs including pancreas, breast tissue, and neurons. MALAT1 was also found to be upregulated in many cancerous tissues, and the expression of MALAT1 strongly correlated with tumor development and metastasis. [12-14]. This makes MALAT1 a very promising target for designing therapeutics to regulate expression levels.

The nuclear localization of MALAT1 restricts its primary functions to nuclear pathways. MALAT1 is found to be localized in nuclear speckles, which are regions in the interchromatin nuclear space that contain a high local concentration of splicing factors and pre-mRNA [15-17]. Nuclear speckles are dynamic structure with constant turnover, and their localization to active transcription sites can be temporal. As such, the primary known function for MALAT1 is splicing regulation and involvement in alternative splicing. MALAT1 can directly bind many

splicing factor proteins and alternative splicing RNA-seq data has revealed thousands of genes whose splicing patterns are regulated by MALAT1 expression [15, 18-20].

Understanding the direct protein interactors to MALAT1 can help reveal information about its function in the nucleus. The McHugh lab has previously identified direct binders of MALAT1 through RNA-Antisense Purification coupled with Mass Spectroscopy (RAP-MS). UV_{254nm} cross-linking of cells fixes direct RNA-protein interactions, and the MALAT1 interactome can be directly isolated through a pooled ssDNA probe mix covering the anti-sense transcript length of MALAT1. The recovered proteins can then be analyzed through mass spectrometry for identification.

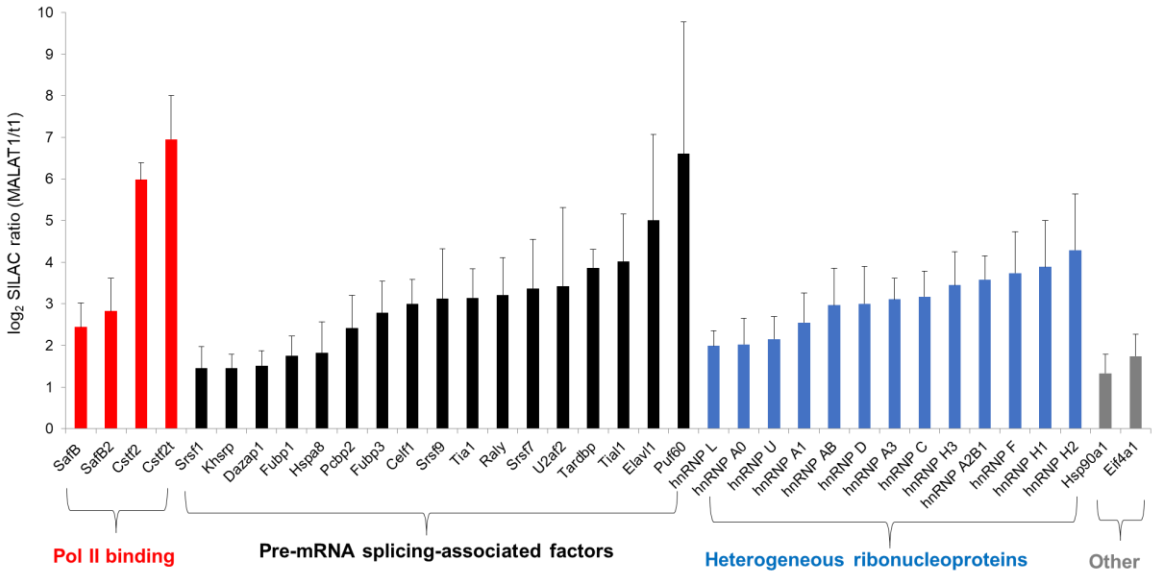


Figure 1.3 - RAP-MS reveals direct protein binders of MALAT1

Unpublished data from the McHugh lab shows a wide range of direct protein binders to lncRNA MALAT1.

RAP-MS reveals three main distinct classes of protein interactors to MALAT1 – Polymerase II binders, pre-mRNA splicing associated factors, and hnRNPs [Figure 1.3]. These reveal the common functions of MALAT1 – pre-mRNA binding and splicing regulation,

transcriptional regulation and mRNA processing, and nuclear export of mRNA [10, 21-23]. Hence, studying the MALAT1/protein direct interactions in disease conditions can help reveal information about MALAT1's role in regulating biological pathways for disease progression or inhibition.

1.3 TDP-43

One of the direct protein binding partners of MALAT1 is the transactive response DNA binding protein (TDP-43). TDP-43 is a commonly expressed protein of 43 kDa and is involved in many biological processes. TDP-43 is encoded by the *TARDBP* gene on chromosome 1 and is localized both within the nucleus and the cytoplasm [24, 25]. The inter-cellular trafficking of TDP-43 is achieved through a nuclear localization signal (NLS) and nuclear export sequence (NES) domains located on the protein. In addition, the protein contains two tandem RNA recognition motifs (RRMs) regions that are recognized as nucleic acid binding domains, and has been structurally solved in complex with UG-rich RNA using NMR [26]. This is followed by a glycine-rich C terminal domain that is intrinsically disordered and can recruit to and bind hnRNP proteins [26-28] [Figure 1.4A].

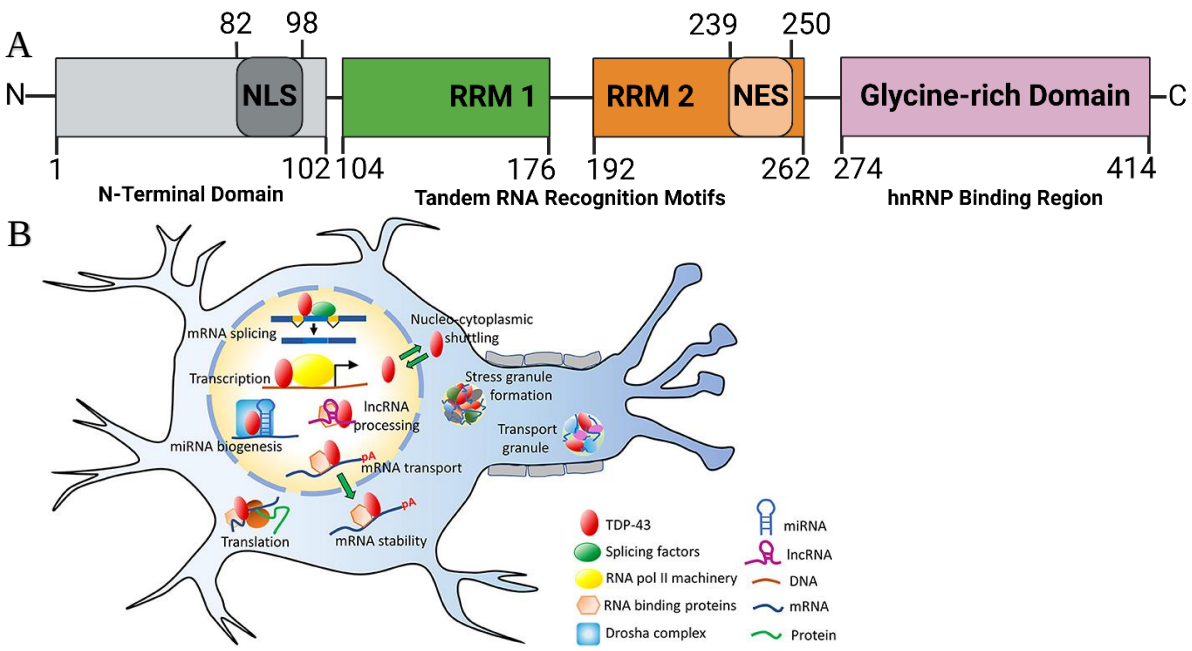


Figure 1.4 - Tar DNA Binding Protein structure and function

(A) The amino acid sequence of TDP-43 organized in domains. (B) An overview of TDP-43 common molecular functions in both nucleus and cytoplasm. This figure is adapted from Frontiers in Molecular Neuroscience [29].

TDP-43 function is commonly determined through its interaction with various mRNA sequences in the nucleus and cytoplasm to regulate gene expression, modulate alternative splicing, and coordinate mRNA trafficking and transport across the nuclear membrane [30-32] [Figure 1.4B]. Cross-linking immunoprecipitation (CLIP) sequencing studies revealed a direct binding interaction between TDP-43 RRMs and long UG-rich RNA regions [34-36]. Minimal binding study experiments have indicated a stoichiometric maximum of 6 UG repeats interacting with one TDP-43 tandem RRM construct [37]. The location of these UG regions vary across the gene, but the re-analysis of CLIP-seq datasets indicate a preference for TDP-43 binding to intronic and 3'UTR UG-rich regions [34-36].

TDP-43 binds a variety of RNA molecules including lncRNAs, premRNAs, and miRNAs through UG repeat regions in a cooperative manner [37]. The solubility of TDP-43 is strongly associated with its binding interaction to ssDNA or RNA targets [38-40] and mutations in the RRM regions that inhibit RNA binding result in TDP-43 aggregates with inhibited functional activity [41, 42]. Dysregulated and aggregated TDP-43 is a common hallmark of many neurodegenerative diseases [29, 30] and an improved understanding to key RNA targets that mediate TDP-43 stability and function seems critical for understanding the importance of TDP-43 toxic inclusions and their role in initiating neurodegeneration.

1.4 MALAT1, TDP-43, and Parkinson's Disease

TDP-43 has been previously identified as an interactor of MALAT1 [43-45]. Earlier studies indicated TDP-43 as a regulator of MALAT1, where the dysregulation of TDP-43 expression levels resulted in dysregulation of MALAT1 in lung cancer cells. Fluorescence microscopy readouts also indicated co-localization of MALAT1 and nuclear TDP-43 in HC11 mammary epithelium cells [43]. CLIP-seq experiments performed with TDP-43 revealed a large region of around 300 nucleotides on MALAT1 from nucleotide position 6638 to 6938 as the binding region, and this binding interaction is conserved across multiple cell lines [34-36]. A unique feature of this binding interaction is the absence of large UG rich regions, commonly identified as the TDP-43 binding motif. Non-canonical TDP-43 binding targets are few – the other common non-canonical binding is to the 3'UTR of the TARDBP gene, where self-binding promotes autoregulation and is a key method in homeostatic control of TDP-43 levels [46]. Thus, the unique nature of the MALAT1/TDP-43 binding may suggest a unique non-canonical function for this interaction.

In addition, CLIP-seq studies performed on diseased brain tissue reveal an increase in the MALAT1-TDP-43 binding interaction compared to healthy normal brain tissue. The MALAT1 binding results in one of the highest fold changes amongst all target transcripts [34]. Coupling this increase in binding in neurodegeneration conditions with the non-canonical interaction with MALAT1 provides an interest in studying the function of the MALAT1/TDP-43 interaction and how it may be relevant in neurodegenerative conditions.

As previously mentioned, the aberrant regulation of TDP-43 is a common hallmark in neurodegeneration that has been widely studied for many years. Its role in the development of Parkinson's Disease (PD) is less studied compared to other common forms of neurodegeneration – namely Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Lobar Degeneration (FTLD) and others [47-51]. In recent studies, patients have been observed with movement disorders and symptoms that closely align to PD, and postmortem neuropathological examination revealed aggregated and overexpressed TDP-43 in diseased brain tissue [52]**.**

Unlike TDP-43, there have been multiple studies on the role of MALAT1 in PD progression. Elevated MALAT1 levels have been observed in a 1-methyl-4-phenyl-1, 2, 3, 6-tetrahydropyridine (MPTP)-induced mice model of PD and methyl-4-phenylpyridinium (MPP +) induced human cell tissue models. [53, 54]. In addition, a high throughput screening assay was performed on PD brain tissue compared to healthy brain controls, and several lncRNAs including MALAT1 exhibited a significant increase in expression levels in the patient samples [53] Mechanistic studies have been sought after to determine the role of MALAT1 in PD progression and the current literature has focused on three different pathways of regulation - perturbed α-synuclein hemostasis [55-57], apoptosis and autophagy [54, 58, 59] and neuro-inflammation [60-62]. Despite the overlapping functional roles between MALAT1 and TDP-43, and the shared

disease relevance and binding interactions, there are no comprehensive studies on the potential role of the MALAT1/TDP-43 interaction on PD progression, and this thesis serves to provide some PD-relevant pathways that may be regulated concurrently by these biomolecules.

1.5 Alternative Splicing and SAT1

Messenger RNA (mRNA) undergo multiple post-transcriptional processing stages that are highly regulated. This ensures that the proper mRNA sequence is present for translation and synthesis of proteins, and disruption of these processes can often lead to disease progression [63, 64]. Once of the key post-transcriptional processes for mRNA is alternative splicing (AS). AS is a mechanism that can induce complexity in higher order organisms, due to its ability to localize multiple gene isoforms within the same locus [65]. The process of alternative splicing involves the selection of specific exons to be translated into proteins, with unselected exons being spliced out of the final mRNA product. This phenomenon is regulated by the spliceosome, a large macrocomplex structure of splicing proteins and small nuclear RNA which help catalyse the splicing of unprocessed pre-mRNA [66, 67].

The selection of exons to be spliced out can be governed by splicing repressors or activators. These are biomolecules that can promote or hinder the selection of specific sites for AS, and thus promote the selection of specific exons in the final isoform product [68]. These biomolecules often tend to be regulatory proteins, although current literature has also revealed a role of non-coding RNA in splice site selection processes in human and other cellular organisms [69, 70].

The favoring of certain isoforms through AS can help alter cell viability in disease conditions, as AS can be a method of transcript regulation at homeostatic levels. This often occurs when one of the isoforms results in the formation of a truncated protein product that exhibits cellular toxicity due to its aberrant function [71]. Such isoforms are often not processed

through translation but are instead degraded in their mRNA form, through a process known as nonsense-mediated decay (NMD) [72, 73]. AS-NMD is a regulatory phenomenon through which many cell-viability regulating genes can be maintained at a steady-state level [74].

An example of a gene regulated by AS-NMD with potential therapeutic effects in a disease condition is the Spermine/Spermidine Acetyltransferase 1 (SAT1) gene [75, 76]. Located on the X chromosome, SAT1 codes for an enzyme that catalyses the addition of an acetyl group onto spermine and spermidine, a subset of polyamines [76]. Polyamines are organic polycations that are positively charged, and have been reported to be directly correlated with cell survivability and growth [77, 78]. The steady-state regulation of polyamines ensures a balance is created to mitigate neuroprotective effects with neurodegenerative effects, which may play a role in the progression of Parkinson's Disease (PD). Studies have shown that polyamine levels tend to be altered in PD-patient brain tissue, and N^1-acetylspermidine, N^8-acetylspermidine, and N^1,N^8-diacetylspermidine have been shown to correlate with cognitive decline as per the Hoehn and Yahr scale [79-81].

Acetylated versions of spermine and spermidine are directly correlated with SAT1 levels, which supports the relevance of SAT1 homeostatic levels for cell survivability [80]. Thus, the methods of SAT1 regulation are important to study, as this can be a focus for potential therapeutics for combating neurodegeneration. SAT1 is known to be regulated via two mechanisms – firstly, the increase in polyamine concentrations results in polyamine binding to upstream polyamine responsive element regions to promote overall transcriptional processivity [82, 83]. Secondly, SAT1 can undergo AS-NMD to produce a nonfunctional transcript through inclusion of an internal 'X' codon, that results in an insertion of a premature stop codon [76,84]. This transcript undergoes NMD and is degraded, thus reducing overall SAT1 levels and

maintaining homeostatic balance. While there have been a few studies on the mechanism through which SAT1 AS-NMD occurs [84, 85], the potential involvement of MALAT1 and TDP-43 in regulating SAT1 AS is a novel direction to pursue and can have significant biological relevance in the context of neurodegenerative disease regulation.

1.6 References

1. Crick, F. (1970). Central dogma of molecular biology. Nature, 227(5258), 561-563.

2. Jarroux, J., Morillon, A., & Pinskaya, M. (2017). History, discovery, and classification of lncRNAs. Long non coding RNA biology, 1-46.

3. Snyder, M. P., Gingeras, T. R., Moore, J. E., Weng, Z., Gerstein, M. B., Ren, B., ... & Myers, R. M. (2020). Perspectives on ENCODE. Nature, 583(7818), 693-698.

4. Marchese, F. P., Raimondi, I., & Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. Genome biology, 18, 1-13.

5. Clemson, C. M., Hutchinson, J. N., Sara, S. A., Ensminger, A. W., Fox, A. H., Chess, A., & Lawrence, J. B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Molecular cell, 33(6), 717–726. https://doi.org/10.1016/j.molcel.2009.01.026

6. Ma, B., Wang, S., Wu, W., Shan, P., Chen, Y., Meng, J., ... & Guo, Y. (2023). Mechanisms of circRNA/lncRNA-miRNA interactions and applications in disease and drug research. Biomedicine & Pharmacotherapy, 162, 114672.

7. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012 Sep;22(9):1775-89. doi: 10.1101/gr.132159.111. PMID: 22955988; PMCID: PMC3431493.

8. Ziats, M. N., and Rennert, O. M. (2013). Aberrant expression of long noncoding RNAs in autistic brain. J. Mol. Neurosci. 49, 589–593. doi: 10.1007/s12031-012-9880–9888

9. Roberts, T. C., Morris, K. V., & Wood, M. J. (2014). The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1652), 20130507.

10. Zhang, X., Hamblin, M. H., & Yin, K. J. (2017). The long noncoding RNA Malat1: Its physiological and pathophysiological functions. RNA Biology, 14(12), 1705–1714. https://doi.org/10.1080/15476286.2017.1358347

11. Ji P, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H, Müller-Tidow C. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene. 2003 Sep 11;22(39):8031-41. doi: 10.1038/sj.onc.1206928. PMID: 12970751.

12. Guffanti, A., Iacono, M., Pelucchi, P., Kim, N., Soldà, G., Croft, L. J., ... & De Bellis, G. (2009). A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. BMC genomics, 10, 1-17.

13. Koshimizu, T. A., Fujiwara, Y., Sakai, N., Shibata, K., & Tsuchiya, H. (2010). Oxytocin stimulates expression of a noncoding RNA tumor marker in a human neuroblastoma cell line. Life sciences, 86(11-12), 455-460.

14. Qi, P., & Du, X. (2013). The long non-coding RNAs, a new cancer diagnostic and therapeutic gold mine. Modern pathology, 26(2), 155-165.

15. Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A. and Blencowe, B.J., 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell, 39(6), pp.925-938.

16. Miyagawa, R., Tano, K., Mizuno, R.I.E., Nakamura, Y.O., Ijiri, K., Rakwal, R., Shibato, J., Masuo, Y., Mayeda, A., Hirose, T. and Akimitsu, N., 2012. Identification of cis-and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. Rna, 18(4), pp.738-751.

17. Chen, Yu, and Andrew S. Belmont. "Genome organization around nuclear speckles." Current opinion in genetics & development 55 (2019): 91-99.

18. Gordon, Michael A., Beatrice Babbs, Dawn R. Cochrane, Benjamin G. Bitler, and Jennifer K. Richer. "The long non-coding RNA MALAT1 promotes ovarian cancer progression by regulating RBFOX2-mediated alternative splicing." Molecular carcinogenesis 58, no. 2 (2019): 196-205.

19. Ghosh, Arpita, Satya Prakash Pandey, Asgar Hussain Ansari, Jennifer Seematti Sundar, Praveen Singh, Yasmeen Khan, Mary Krishna Ekka, Debojyoti Chakraborty, and Souvik Maiti. "Alternative splicing modulation mediated by G-quadruplex structures in MALAT1 lncRNA." Nucleic Acids Research 50, no. 1 (2022): 378.

20. Miao, H., Wu, F., Li, Y., Qin, C., Zhao, Y., Xie, M., Dai, H., Yao, H., Cai, H., Wang, Q. and Song, X., 2022. MALAT1 modulates alternative splicing by cooperating with the splicing factors PTBP1 and PSF. Science Advances, 8(51), p.eabq7289.

21. Arun, Gayatri, Disha Aggarwal, and David L. Spector. "MALAT1 long non-coding RNA: functional implications." Non-coding RNA 6, no. 2 (2020): 22.

22. Yoshimoto, Rei, Akila Mayeda, Minoru Yoshida, and Shinichi Nakagawa. "MALAT1 long non-coding RNA in cancer." Biochimica et biophysica acta (BBA)-gene regulatory mechanisms 1859, no. 1 (2016): 192-199.

23. Gutschner, Tony, Monika Hämmerle, and Sven Diederichs. "MALAT1—a paradigm for long noncoding RNA function in cancer." Journal of molecular medicine 91 (2013): 791-801.

24. Chen-Plotkin, Alice S., Virginia M-Y. Lee, and John Q. Trojanowski. "TAR DNA-binding protein 43 in neurodegenerative disease." Nature Reviews Neurology 6, no. 4 (2010): 211-220.

25. Chou, C.C., Zhang, Y.I., Umoh, M.E., Vaughan, S.W., Lorenzini, I., Liu, F., Sayegh, M., Donlin-Asp, P.G., Chen, Y.H., Duong, D.M. and Seyfried, N.T., 2018. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. Nature neuroscience, 21(2), pp.228-239.

26. Lukavsky, Peter J., Dalia Daujotyte, James R. Tollervey, Jernej Ule, Cristiana Stuani, Emanuele Buratti, Francisco E. Baralle, Fred F. Damberger, and Frédéric HT Allain. "Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43." Nature structural & molecular biology 20, no. 12 (2013): 1443-1449.

27. Wei, Yuanyuan, Liangzhong Lim, Lu Wang, and Jianxing Song. "Inter-domain interactions of TDP-43 as decoded by NMR." Biochemical and biophysical research communications 473, no. 2 (2016): 614-619.

28. Mompeán, Miguel, Valentina Romano, David Pantoja-Uceda, Cristiana Stuani, Francisco E. Baralle, Emanuele Buratti, and Douglas V. Laurents. "The TDP-43 N-terminal domain structure at high resolution." The FEBS journal 283, no. 7 (2016): 1242-1260.

29. Prasad, Archana, Vidhya Bharathi, Vishwanath Sivalingam, Amandeep Girdhar, and Basant K. Patel. "Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis." Frontiers in molecular neuroscience 12 (2019): 25.

30. Cohen, Todd J., Virginia MY Lee, and John Q. Trojanowski. "TDP-43 functions and pathogenic mechanisms implicated in TDP-43 proteinopathies." Trends in molecular medicine 17, no. 11 (2011): 659-667.

31. Nagano, S., Jinno, J., Abdelhamid, R.F., Jin, Y., Shibata, M., Watanabe, S., Hirokawa, S., Nishizawa, M., Sakimura, K., Onodera, O. and Okada, H., 2020. TDP-43 transports ribosomal

protein mRNA to regulate axonal local translation in neuronal axons. Acta Neuropathologica, 140, pp.695-713.

32. Štalekar, Maja, Xiaoke Yin, Katja Rebolj, Simona Darovic, Claire Troakes, Manuel Mayr, Christopher E. Shaw, and Boris Rogelj. "Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport." Neuroscience 293 (2015): 157-170.

33. Ma, X.R., Prudencio, M., Koike, Y., Vatsavayai, S.C., Kim, G., Harbinski, F., Briner, A., Rodriguez, C.M., Guo, C., Akiyama, T. and Schmidt, H.B., 2022. TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A. Nature, 603(7899), pp.124-130.

34. Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Župunski, V. and Patani, R., 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nature neuroscience, 14(4), pp.452-458.

35. Wu, Huang, Qing-Fei Yin, Zheng Luo, Run-Wen Yao, Chuan-Chuan Zheng, Jun Zhang, Jian-Feng Xiang, Li Yang, and Ling-Ling Chen. "Unusual processing generates SPA LncRNAs that sequester multiple RNA binding proteins." Molecular cell 64, no. 3 (2016): 534-548.

36. Hallegger, M., Chakrabarti, A.M., Lee, F.C., Lee, B.L., Amalietti, A.G., Odeh, H.M., Copley, K.E., Rubien, J.D., Portz, B., Kuret, K. and Huppertz, I., 2021. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell, 184(18), pp.4680-4696.

37. Rengifo-Gonzalez, Juan Carlos, Krystel El Hage, Marie-Jeanne Clément, Emilie Steiner, Vandana Joshi, Pierrick Craveur, Dominique Durand, David Pastré, and Ahmed Bouhss. "The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation." Elife 10 (2021): e67605.

38. Pérez-Berlanga, M., Wiersma, V.I., Zbinden, A., De Vos, L., Wagner, U., Foglieni, C., Mallona, I., Betz, K.M., Cléry, A., Weber, J. and Guo, Z., 2023. Loss of TDP-43 oligomerization or RNA binding elicits distinct aggregation patterns. The EMBO journal, 42(17), p.e111719.

39. Huang, Yi-Chen, Ku-Feng Lin, Ruei-Yu He, Pang-Hsien Tu, Jiri Koubek, Yin-Chih Hsu, and Joseph Jen-Tse Huang. "Inhibition of TDP-43 aggregation by nucleic acid binding." PloS one 8, no. 5 (2013): e64002.

40. Mann Mann, J.R., Gleixner, A.M., Mauna, J.C., Gomes, E., DeChellis-Marks, M.R., Needham, P.G., Copley, K.E., Hurtle, B., Portz, B., Pyles, N.J. and Guo, L., 2019. RNA binding antagonizes neurotoxic phase transitions of TDP-43. Neuron, 102(2), pp.321-338.

41. Chen, H.J., Topp, S.D., Hui, H.S., Zacco, E., Katarya, M., McLoughlin, C., King, A., Smith, B.N., Troakes, C., Pastore, A. and Shaw, C.E., 2019. RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. Brain, 142(12), pp.3753-3770.

42. Sun, Han, Wei Chen, Lin Chen, and Wenqing Zheng. "Exploring the molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43 using molecular dynamics simulation and free energy calculation." Journal of Computational Chemistry 42, no. 23 (2021): 1670-1680.

43. Nguyen, T.M., Kabotyanski, E.B., Reineke, L.C., Shao, J., Xiong, F., Lee, J.H., Dubrulle, J., Johnson, H., Stossi, F., Tsoi, P.S. and Choi, K.J., 2020. The SINEB1 element in the long non-coding RNA Malat1 is necessary for TDP-43 proteostasis. Nucleic acids research, 48(5), pp.2621-2642.

44. Guo, Fengjie, Feng Jiao, Zuoqing Song, Shujun Li, Bin Liu, Hongwei Yang, Qinghua Zhou, and Zhigang Li. "Regulation of MALAT1 expression by TDP43 controls the migration and invasion of non-small cell lung cancer cells in vitro." Biochemical and biophysical research communications 465, no. 2 (2015): 293-298.

45. Liu, W., Wang, Z., Liu, L., Yang, Z., Liu, S., Ma, Z., Liu, Y., Ma, Y., Zhang, L., Zhang, X. and Jiang, M., 2020. LncRNA Malat1 inhibition of TDP43 cleavage suppresses IRF3-initiated antiviral innate immunity. Proceedings of the National Academy of Sciences, 117(38), pp.23695-23706.

46. Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E. and Baralle, F.E., 2011. TDP-43 regulates its mRNA levels through a negative feedback loop. The EMBO journal, 30(2), pp.277-288.

47. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun. 2006 Dec 22;351(3):602-11. doi: 10.1016/j.bbrc.2006.10.093. Epub 2006 Oct 30. PMID: 17084815.

48. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006 Oct 6;314(5796):130-3. doi: 10.1126/science.1134108. PMID: 17023659.

49. Hasegawa M, Arai T, Akiyama H, Nonaka T, Mori H, Hashimoto T, Yamazaki M, Oyanagi K. TDP-43 is deposited in the Guam parkinsonism-dementia complex brains. Brain. 2007 May;130(Pt 5):1386-94. doi: 10.1093/brain/awm065. Epub 2007 Apr 17. PMID: 17439983.

50. Nakashima-Yasuda H, Uryu K, Robinson J, Xie SX, Hurtig H, Duda JE, Arnold SE, Siderowf A, Grossman M, Leverenz JB, Woltjer R, Lopez OL, Hamilton R, Tsuang DW, Galasko D, Masliah E, Kaye J, Clark CM, Montine TJ, Lee VM, Trojanowski JQ. Co-morbidity of TDP-43 proteinopathy in Lewy body related diseases. Acta Neuropathol. 2007 Sep;114(3):221-9. doi: 10.1007/s00401-007-0261-2. Epub 2007 Jul 25. PMID: 17653732.

51. Nelson, P.T., Dickson, D.W., Trojanowski, J.Q., Jack, C.R., Boyle, P.A., Arfanakis, K., Rademakers, R., Alafuzoff, I., Attems, J., Brayne, C. and Coyle-Gilchrist, I.T., 2019. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. Brain, 142(6), pp.1503-1527.

52. Yamashita, Rika, Goichi Beck, Yuki Yonenobu, Kimiko Inoue, Akihiko Mitsutake, Hiroyuki Ishiura, Masato Hasegawa, Shigeo Murayama, and Hideki Mochizuki. "TDP-43 proteinopathy presenting with typical symptoms of Parkinson's disease." Mov Disord 37, no. 7 (2022): 1561-1563.

53. Kraus, Theo FJ, Melanie Haider, Judith Spanner, Martina Steinmaurer, Vanessa Dietinger, and Hans A. Kretzschmar. "Altered long noncoding RNA expression precedes the course of Parkinson's disease—a preliminary report." Molecular neurobiology 54, no. 4 (2017): 2869-2877.

54. Lu, Yi, Zhongying Gong, Xiaojie Jin, Peng Zhao, Yuting Zhang, and Zhiyun Wang. "LncRNA MALAT1 targeting miR-124-3p regulates DAPK1 expression contributes to cell apoptosis in Parkinson's Disease." Journal of cellular biochemistry 121, no. 12 (2020): 4838-4848.

55. Dehay, B., Bourdenx, M., Gorry, P., Przedborski, S., Vila, M., Hunot, S., Singleton, A., Olanow, C.W., Merchant, K.M., Bezard, E. and Petsko, G.A., 2015. Targeting α-synuclein for treatment of Parkinson's disease: mechanistic and therapeutic considerations. The Lancet Neurology, 14(8), pp.855-866.

56. Bennett MC. The role of alpha-synuclein in neurodegenerative diseases. Pharmacol Ther. 2005 Mar;105(3):311-31. doi: 10.1016/j.pharmthera.2004.10.010. Epub 2004 Dec 8. PMID: 15737408.

57. Zhang QS, Wang ZH, Zhang JL, Duan YL, Li GF, Zheng DL. Beta-asarone protects against MPTP-induced Parkinson's disease via regulating long non-coding RNA MALAT1 and inhibiting α-synuclein protein expression. Biomed Pharmacother. 2016 Oct;83:153-159. doi: 10.1016/j.biopha.2016.06.017. Epub 2016 Jun 27. PMID: 27470562.

58. Chen, Qin, Xiaoyan Huang, and Renjie Li. "lncRNA MALAT1/miR-205-5p axis regulates MPP+-induced cell apoptosis in MN9D cells by directly targeting LRRK2." American Journal of Translational Research 10, no. 2 (2018): 563.

59. Lv, Kefeng, Yuhua Liu, Yanbing Zheng, Shaowen Dai, Peifeng Yin, and Haifeng Miao. "Long non-coding RNA MALAT1 regulates cell proliferation and apoptosis via miR-135b-5p/GPNMB axis in Parkinson's disease cell model." Biological Research 54 (2021).

60. Puthanveetil P, Chen S, Feng B, Gautam A, Chakrabarti S. Long non-coding RNA MALAT1 regulates hyperglycaemia induced inflammatory process in the endothelial cells. J Cell Mol Med. 2015 Jun;19(6):1418-25. doi: 10.1111/jcmm.12576. Epub 2015 Mar 19. PMID: 25787249; PMCID: PMC4459855.

61. Ding Y, Guo F, Zhu T, Li J, Gu D, Jiang W, Lu Y, Zhou D. Mechanism of long non-coding RNA MALAT1 in lipopolysaccharide-induced acute kidney injury is mediated by the miR-146a/NF-κB signaling pathway. Int J Mol Med. 2018 Jan;41(1):446-454. doi: 10.3892/ijmm.2017.3232. Epub 2017 Nov 2. PMID: 29115409.

62. Cai, Li-Jun, Li Tu, Xiao-Mo Huang, Jia Huang, Nan Qiu, Guang-Hong Xie, Jian-Xiong Liao, Wei Du, Ying-Yue Zhang, and Jin-Yong Tian. "LncRNA MALAT1 facilitates inflammasome activation via epigenetic suppression of Nrf2 in Parkinson's disease." Molecular brain 13 (2020): 1-15.

63. Corbett A. H. (2018). Post-transcriptional regulation of gene expression and human disease. Current opinion in cell biology, 52, 96–104. https://doi.org/10.1016/j.ceb.2018.02.011

64. Carey, K. T., & Wickramasinghe, V. O. (2018). Regulatory potential of the RNA processing machinery: implications for human disease. Trends in Genetics, 34(4), 279-290.

65. Tao, Yining, Qi Zhang, Haoyu Wang, Xiyu Yang, and Haoran Mu. "Alternative splicing and related RNA binding proteins in human health and disease." Signal Transduction and Targeted Therapy 9, no. 1 (2024): 26.

66. Yan, C., Wan, R., & Shi, Y. (2019). Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. Cold Spring Harbor perspectives in biology, 11(1), a032409.

67. Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. cell, 136(4), 701-718.

68. Ule, J., & Blencowe, B. J. (2019). Alternative splicing regulatory networks: functions, mechanisms, and evolution. Molecular cell, 76(2), 329-345.

69. Pisignano, G., & Ladomery, M. (2021). Epigenetic regulation of alternative splicing: how LncRNAs tailor the message. Non-coding RNA, 7(1), 21.

70. Bardou, Florian, Federico Ariel, Craig G. Simpson, Natali Romero-Barrios, Philippe Laporte, Sandrine Balzergue, John WS Brown, and Martin Crespi. "Long noncoding RNA modulates alternative splicing regulators in Arabidopsis." Developmental cell 30, no. 2 (2014): 166-176.

71. Nikom, D., & Zheng, S. (2023). Alternative splicing in neurodegenerative disease and the promise of RNA therapies. Nature Reviews Neuroscience, 24(8), 457-473.

72. Brogna, S., & Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nature structural & molecular biology, 16(2), 107-113.

73. Tan, K., Stupack, D. G., & Wilkinson, M. F. (2022). Nonsense-mediated RNA decay: an emerging modulator of malignancy. Nature Reviews Cancer, 22(8), 437-451.

74. Neff, A. T., Lee, J. Y., Wilusz, J., Tian, B., & Wilusz, C. J. (2012). Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome research, 22(8), 1457-1467.

75. Pegg, A. E. (2009). Mammalian polyamine metabolism and function. IUBMB life, 61(9), 880-894.

76. Pegg, A. E. (2008). Spermidine/spermine-N 1-acetyltransferase: a key metabolic regulator. American Journal of Physiology-Endocrinology and Metabolism, 294(6), E995-E1010.

77. Mandal, S., Mandal, A., Johansson, H. E., Orjalo, A. V., & Park, M. H. (2013). Depletion of cellular polyamines, spermidine and spermine, causes a total arrest in translation and growth in mammalian cells. Proceedings of the National Academy of Sciences, 110(6), 2169-2174.

78. Mandal, S., Mandal, A., & Park, M. H. (2015). Depletion of the polyamines spermidine and spermine by overexpression of spermidine/spermine N 1-acetyltransferase 1 (SAT1) leads to mitochondria-mediated apoptosis in mammalian cells. Biochemical Journal, 468(3), 435-447.

79. Saiki, S., Sasazawa, Y., Fujimaki, M., Kamagata, K., Kaga, N., Taka, H., Li, Y., Souma, S., Hatano, T., Imamichi, Y., Furuya, N., Mori, A., Oji, Y., Ueno, S. I., Nojiri, S., Miura, Y., Ueno, T., Funayama, M., Aoki, S., & Hattori, N. (2019). A metabolic profile of polyamines in parkinson disease: A promising biomarker. Annals of neurology, 86(2), 251–263. https://doi.org/10.1002/ana.25516

80. Vrijsen, S., Houdou, M., Cascalho, A., Eggermont, J., & Vangheluwe, P. (2023). Polyamines in Parkinson's disease: balancing between neurotoxicity and neuroprotection. Annual Review of Biochemistry, 92, 435-464.

81. Lewandowski, N.M., Ju, S., Verbitsky, M., Ross, B., Geddie, M.L., Rockenstein, E., Adame, A., Muhammad, A., Vonsattel, J.P., Ringe, D. and Cote, L., 2010. Polyamine pathway contributes to the pathogenesis of Parkinson disease. Proceedings of the National Academy of Sciences, 107(39), pp.16970-16975.

82. Wang, Y., Xiao, L., Thiagalingam, A., Nelkin, B. D., & Casero, R. A. (1998). The identification of a cis-element and a trans-acting factor involved in the response to polyamines and polyamine analogues in the regulation of the human spermidine/spermine N 1-acetyltransferase gene transcription. Journal of Biological Chemistry, 273(51), 34623-34630.

83. Wang, Y., & Casero Jr, R. A. (2006). Mammalian polyamine catabolism: a therapeutic target, a pathological problem, or both?. Journal of biochemistry, 139(1), 17-25.

84. Baker, Mai, Mayra Petasny, Nadeen Taqatqa, Mercedes Bentata, Gillian Kay, Eden Engal, Yuval Nevo, Ahmad Siam, Sara Dahan, and Maayan Salton. "KDM3A regulates alternative splicing of cell-cycle genes following DNA damage." Rna 27, no. 11 (2021): 1353-1362.

85. Pozzi, B., Bragado, L., Mammi, P., Torti, M.F., Gaioli, N., Gebhard, L.G., García Solá, M.E., Vaz-Drago, R., Iglesias, N.G., García, C.C. and Gamarnik, A.V., 2020. Dengue virus targets RBM10 deregulating host cell splicing and innate immune response. Nucleic acids research, 48(12), pp.6824-6838.

Chapter 2 : MALAT1 RNA levels affect cell viability and modulate TDP-43 binding to mRNA in the nucleus

Abstract

TAR DNA-binding protein (TDP-43) and Metastasis Associated Lung Adenocarcinoma Transcript (MALAT1) RNA are two abundantly expressed macromolecules in the human cell nucleus. Increased interaction of TDP-43 and MALAT1, as well as dysregulation of TDP-43 function, was previously identified in brain samples from patients with neurodegenerative disease compared to healthy brain tissues. We hypothesized that TDP-43 function may depend in part on MALAT1 expression levels, due to the coordinated regulation and interaction between these two macromolecules that has been observed in multiple cell types. Here, we show that alterations in MALAT1 expression affect cell viability and can modulate TDP-43 binding to other messenger RNAs in HEK293 and SH-SH5Y human cell lines. Depletion of MALAT1 RNA enabled increased binding of TDP-43 to messenger RNA transcripts, suggestive of a redistribution of TDP-43 protein binding. Disruption of either MALAT1 or TDP-43 expression induced cell death, indicating that both macromolecules contribute positively to survival in SH SY5Y neuroblastoma cells. Finally, we examined the contribution of MALAT1 expression to survival in a cell culture model of neurodegeneration using MPP^+ treatment in SH-SY5Y cells. Depletion of MALAT1 RNA protects against toxicity in a cellular model of neurodegeneration and modulates TDP-43 binding to mRNA transcripts involved in apoptotic cell death. These data suggest that coordinated regulation of a network of RNA and protein interactions could provide a mechanism to maintain appropriate non-coding RNA and messenger RNA expression levels in healthy human cells. Dysregulation of the MALAT1 RNA may affect human cell viability in early neurodegenerative disease processes.

Introduction

RNA-protein interactions regulate gene expression, splicing, and cell growth. Genome-wide studies of RNA-protein interactions suggest that eukaryotic cells have a complex network of binding interactions between RNAs and proteins, including mRNAs, non-coding RNAs, splicing factors and other RNA binding proteins. TAR DNA-binding protein 43 (TARDBP or TDP-43) is a ubiquitously expressed multifunctional nucleic acid binding protein which has been associated with transcriptional regulation, RNA stability, and alternative splicing of mRNAs in human cells. Dysfunction of TDP-43 is frequently observed in human neurodegenerative disease, particularly in brain tissues from patients with amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD), Parkinson's disease (PD), and limbic-predominant age-related TDP-43 encephalopathy (LATE) [1, 2, 3, 4, 5]. The underlying molecular interactions which result in TDP-43 dysfunction – and whether these interactions are a cause or consequence of human neurodegeneration - remain unknown.

TDP-43 may act in modulating RNA and protein networks in healthy and diseased cells by affecting alternative splicing and gene expression. TDP-43 binds thousands of RNA transcripts in mammalian cells, based on crosslinking and immunoprecipitation experiments with high-throughput sequencing, including CLIP-seq and iCLIP studies. TDP-43 binds specifically to messenger RNA and non-coding RNA targets containing sequential UG repeats [6, 7]. Binding to RNA can modulate TDP-43 protein solubility and toxicity in cells [8, 9]. An examination of TDP-43 RNA binding in post-mortem brain samples of patients with FTD identified the metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) RNA as one of the transcripts with the largest increases in TDP-43 binding in diseased tissues [10]. It has not yet been investigated whether MALAT1 RNA levels can affect TDP-43 function in neuronal cells.

MALAT1 is an 8.5 Kb non-coding RNA with proposed roles in mRNA processing, gene regulation, and alternative splicing [11, 12]. MALAT1 is abundantly expressed and predominantly nuclear localized in many human cell types. MALAT1 was first identified in metastatic lung cancer cells [13] and, like TDP-43, has since been implicated in multiple biological processes, including the regulation of gene expression, alternative splicing, and neuronal synapse formation [14]. MALAT1 and TDP-43 expression have been linked in previous studies. MALAT1 RNA levels are altered by perturbation of TDP-43 expression in lung cancer cells [15]. Furthermore, MALAT1 binds directly to TDP-43 and can affect TDP-43 proteostasis [16, 17, 18]. Taken together, these interactions suggest a role for MALAT1 non-coding RNA in modulating TDP-43 function. However, the mechanisms by which MALAT1 RNA could affect TDP-43 activity and potentially contribute to neurodegeneration remain to be fully elucidated. We hypothesized that TDP-43 function may depend on MALAT1 expression levels in human cells and examined whether direct perturbation of MALAT1 RNA levels affected TDP-43 binding to mRNA transcripts related to neuronal survival and function. In this study, we also investigated the consequences of MALAT1 and TDP-43 dysregulation in SH-SY5Y neuroblastoma cells in culture, and in a model of neurodegeneration after drug treatment, to evaluate the importance of this RNA-protein interaction in contributing to human cell survival.

Results

Disruption of either MALAT1 or TDP-43 induces cell death and perturbs the RNA transcript levels of the other binding partner in SH-SY5Y neuroblastoma cells

We evaluated the effects of MALAT1 and TDP-43 expression changes in neuroblastoma cells to determine whether these two molecules contributed to cell survival. TDP-43 and MALAT1 levels were altered to either increase expression through ectopic overexpression of a transgene from a pcDNA vector or decrease expression through RNA knockdown by antisense oligonucleotide (ASO) GapmeR targeting of the specific RNA transcript. Either an increase or decrease in TDP-43 levels led to a significant decrease in SH-SY5Y cell viability (Figure 1A). Next, we measured the levels of MALAT1 after TDP-43 depletion or overexpression, and vice versa, to determine if their expression levels were correlated. Previous research showed that overexpression of TDP-43 stimulated an increase in MALAT1 RNA expression level, while knockdown of TDP-43 caused a decrease in MALAT1 RNA in A549 lung cancer cells [15]. Our experiments revealed that TDP-43 and MALAT1 expression levels are similarly linked in the SH-SY5Y neuroblastoma cell line (Fig. 1B).

We further examined the effects of MALAT1 knockdown or overexpression on cell viability and TDP-43 mRNA levels in SH-SY5Y cells to determine if MALAT1 conversely affects TDP-43 expression or function. An increase in MALAT1 levels did not significantly affect cell viability, while a decrease in MALAT1 expression levels led to cell death in SH-SY5Y neuroblastoma cells (Fig. 1C). Overexpression of MALAT1 caused a significant increase in TDP-43 mRNA level while depletion of MALAT1 RNA did not alter TDP-43 mRNA expression level in SH-SY5Y cells (Fig. 1D). We next evaluated the protein level of TDP-43 by Western blotting. TDP-43 protein levels did not change significantly during the 48 hours after MALAT1 knockdown (Fig. 1E) or MALAT1 overexpression (Fig. 1F).

MALAT1 KD induces cell death in SH-SY5Y cells but does not lead to cytoplasmic localization of TDP-43

To evaluate the effects of decreased MALAT1 levels on cell survival in SH-SY5Y, we used fluorescence microscopy to examine the nuclei of SH-SY5Y cells after 48 hours of GapmeR control or MALAT1-targeting GapmeR ASO transfection. MALAT1 knockdown cells showed disrupted nuclear morphology compared to control cells (Fig. 2A). Classification and quantitation of nuclear area and morphology was performed using the Nuclear Morphology Analysis ImageJ tool [19]. The nuclear area was significantly decreased after MALAT1 knockdown (Fig. 2B), and MALAT1 knockdown cells had a significant increase in the number of apoptotic and irregular nuclei compared to control cells (Fig. 2C). This is consistent with previous observations of nuclear effects after MALAT1 depletion in HeLa cells [12].

In several previous studies, induction of extreme cellular stress or complete loss of RNA binding by TDP-43 resulted in relocalization of the majority of TDP-43 protein to the cytoplasm (reviewed in [20]). We examined the localization of TDP-43 after MALAT1 knockdown through immunofluorescence microscopy. After MALAT1 disruption, most TDP-43 protein remained in the cell nucleus and there was no obvious relocalization of TDP-43 to the cytoplasm or aggregation into nuclear speckles observed in either the SH-SY5Y or HEK293 cell lines (Fig. 2D). Cellular fractionation studies confirmed the microscopy results and indicated that most TDP-43 remained localized in the nucleus after MALAT1 knockdown (Fig. 2E and F).

TDP-43 binds the 3’ UTR of mRNA targets in multiple cell types

We next investigated which RNA transcripts were commonly bound across multiple cell types by TDP-43. Previously published CLIP-seq datasets were reanalyzed to identify targets of

TDP-43 in HEK cells, SH-SY5Y cells and H9 embryonic stem cells [10, 21, 22]. A total of 1549 transcripts were identified as bound by TDP-43 in all three cell type datasets (Fig. 3A). The transcripts in the overlapping TDP-43 bound RNA set were significantly enriched in gene ontology functions associated with Parkinson's Disease, Alzheimer's Disease, Amyotrophic Lateral Sclerosis, oxidative phosphorylation, RNA degradation and pathways of neurodegeneration (Fig. 3B). These gene ontology classifications are consistent with the previously suggested functions of TDP-43 in RNA biology and the progression of neurodegenerative diseases. We next examined where TDP-43 was bound on each mRNA transcript in the commonly bound gene set across the three cell types. TDP-43 has previously been shown to interact at both introns and 3' UTRs of mRNA transcripts [10]. Consistent with previous studies, we identified TDP-43 binding sites at introns, 5' UTRs, and 3' UTRs of mRNA transcripts in HEK, SH-SY5Y, and H9 cells. When normalized by length of each genomic region, we found that the highest proportion of TDP-43 binding occurred in the 3' untranslated region (3' UTR) of the commonly bound mRNA transcript set across the three different cell types (Fig. 3C). TDP-43 binding in the 3' UTR of mRNA transcripts has previously been shown to affect mRNA expression levels by either increasing or decreasing RNA transcript stability [23, 24, 25]. We focused the remaining experiments on mRNA transcripts that were bound by TDP-43 in the 3' UTR in all three cell types, to further investigate the effects of MALAT1 perturbation on TDP-43 binding to this class of mRNA transcripts.

For further mechanistic studies, we selected three mRNA transcripts from the overlapping TDP-43 bound set which were expressed at high, moderate, and low levels in HEK 293 and SH-SY5Y cells: SLC1A5, CSNK1E, and HMGB2. The genes encoding these transcripts are all associated with neuronal survival and function. HMGB2 promotes neural stem cell proliferation

and affects microglial survival [26]. The SLC1A5 gene, also called Alanine, Serine, Cysteine Transporter 2, encodes an amino acid transporter important for neuronal cell function [27]. CSNK1E belongs to the casein kinase family which is necessary for cell cycle progression and neuronal function [28, 29]. CSNK1E has previously been shown to be regulated by TDP-43 in sporadic ALS [30]. The BARD1 mRNA transcript was selected as a negative control since this transcript did not show TDP-43 binding in any of the examined CLIP-seq experiments, even though it is expressed in these cell types. TDP-43 binding sites on mRNA transcripts that were identified based on CLIP-seq peak calling were also examined visually using the Integrative Genomics Viewer (IGV). In the majority of the genes with TDP-43 CLIP binding, a single major peak was identified. For each of the mRNA transcripts SLC1A5, CSNK1E, and HMGB2, an RNA region of three hundred nucleotides in length was identified which encompassed the maximum CLIP-seq read alignment peak on the transcript (Fig. 3D). The CLIP-seq peaks on messenger RNA transcripts in the TDP-43 bound gene set correlated with the presence of sequential UG repeats in the target RNA sequence, consistent with previous findings of $[UG]_n$ as a consensus RNA binding site for TDP-43 [31]. We also identified the primary binding site of TDP-43 on human MALAT1 RNA between nucleotides 6638 to 6938 ($MALAT1_{6638\text{-}6938}$), which was similarly identified as the primary binding site on MALAT1 in the three CLIP-seq datasets examined. The BARD1 negative control transcript has moderate expression in most cell types but did not show TDP-43 CLIP-seq peaks at any location on the transcript in any of the three data sets. Therefore, a randomly selected size-matched region of BARD1 was used as a negative control for TDP-43 binding experiments with mRNA transcripts.

TDP-43 binds the 3' UTR of messenger RNAs with high affinity *in vitro*

To validate the direct RNA-protein interaction regions identified by CLIP-seq analysis, we measured the *in vitro* binding affinities of TDP-43 RRMs for RNA transcript regions from the selected genes. A fluorescent multiplexed electrophoretic mobility shift assay (mEMSA) was used to calculate the equilibrium dissociation constant (K_d) of TDP-43 protein interactions with the different RNA constructs [32]. For this assay, TDP-43 RRM binding was quantified *in vitro* using fluorescently labeled purified transcribed RNA regions from MALAT1 non-coding RNA, the messenger RNA transcripts SLC1A5, CSNK1E, and HMGB2, the positive control TARDBP mRNA region, or the negative control mRNA transcript BARD1 (Fig. 4, Table 1). The K_d for each RNA-protein interaction was quantified based on three to six replicate assays for each pair of binding partners. The positive control TARDBP mRNA 3' UTR region was selected because it was previously identified as a TDP-43 binding site for autoregulation of the transcript at an RNA level [33, 34, 35]. We observed that the measured TDP-43 binding affinity *in vitro* correlated positively with the CLIP-seq signal for the mRNA transcripts (Table 1). The affinity of TDP-43 for each mRNA transcript was also correlated with the number of sequential UG repeats present in the TDP-43 binding region of the messenger RNAs used for *in vitro* binding studies. These positive correlations suggest that *in vitro* TDP-43 binding affinities are likely to be biologically relevant and representative of the *in vivo* interactions between TDP-43 and messenger RNA transcripts captured by CLIP-seq studies. The TARDBP positive control RNA transcript region bound TDP-43 at 272 ± 15 nM. The selected MALAT1 RNA transcript region bound TDP-43 with high affinity at 127 ± 6 nM, in alignment with the observation that MALAT1 is bound by TDP-43 in all datasets. As expected, the negative control BARD1 RNA binding affinity for TDP-43 was significantly lower than any of the specifically interacting TDP-43 mRNA targets, with a calculated K_d of 1790 ± 51 nM.

Reduced level of MALAT1 RNA results in relocalization of TDP-43 binding to messenger RNA transcripts in HEK293 cells

To quantify the effects of MALAT1 RNA perturbation on TDP-43 RNA binding to the 3' UTR of mRNA transcripts in cells, we performed immunoprecipitation of TDP-43 from HEK293 cell lysates, followed by quantitative RT-PCR (IP-qPCR) of selected mRNA targets in cells with or without MALAT1 knockdown treatment [36]. IP-qPCR experiments showed strong enrichment of MALAT1 RNA in TDP-43 captures compared to IgG control captures, and additionally validated TDP-43 binding in cells to the messenger RNAs SLC1A5, CSNK1E, and HMGB2 that were identified in the commonly bound CLIP-seq dataset (Fig. 5A). As expected, the TARDBP positive control RNA transcript was enriched in TDP-43 captures compared to IgG captures, while the BARD1 negative control RNA transcript was not significantly enriched in either the IgG or TDP-43 captures.

Since TDP-43 protein levels were not affected by MALAT1 knockdown and TDP-43 remained predominantly localized to the nucleus, we reasoned that any TDP-43 protein released after MALAT1 RNA knockdown could potentially re-localize to other mRNA targets within the nucleus. Next, we validated the TDP-43 capture antibody efficiency by Western blot analysis of TDP-43 protein elution from immunoprecipitations performed in MALAT1 GapmeR or control GapmeR knockdown cell lysates. In both conditions, equivalent amounts of TDP-43 protein were successfully captured, with no TDP-43 protein present in the negative control IgG pulldowns (Fig. 5B). We repeated the TDP-43 immunoprecipitation studies from nuclear extracts of HEK293 cells after 24 hours of treatment with negative control (NC) GapmeR or MALAT1 GapmeR and compared the binding of TDP-43 to each mRNA transcript between the treatment

conditions. Loss of MALAT1 expression resulted in a significant increase in the binding of nuclear TDP-43 to the SLC1A5, CSNK1E, and HMGB2 transcripts (Fig. 5C). As expected, MALAT1 depletion had no effect on the interaction of TDP-43 with the negative control BARD1 mRNA transcript. After MALAT1 depletion, the overall expression level of each of the target mRNA transcripts CSNK1E, SLC1A5, and HMGB2 increased. Conversely, depletion of TDP-43 using an ASO GapmeR caused a significant decrease in the expression level of these mRNA transcripts (Fig. 5D), suggesting that TDP-43 binding at the 3' UTR stabilizes or promotes an increase in the RNA levels of these transcripts under normal conditions.

Depletion of MALAT1 RNA protects against MPP^+ toxicity in a cell culture model of neurodegeneration

MALAT1 depletion led to alterations in the RNA binding of nuclear TDP-43 and we observed significant cell death after knockdown of MALAT1 in SH-SY5Y cells. Therefore, we tested whether MALAT1 function contributes to cell survival in a model of neurodegeneration by evaluating the effects of alterations in MALAT1 and TDP-43 on cell viability and gene expression in SH-SY5Y neuroblastoma cells exposed to the neurotoxin 1-methyl-4-phenylpyridinium (MPP^+).

Treatment with MPP^+ induces apoptotic cell death in human SH-SY5Y neuroblastoma cells and has been previously used as a cell culture model for neuronal death in neurodegenerative disease [37, 38]. We observed that MALAT1 RNA levels increased significantly in SH-SY5Y cells treated with MPP^+, in line with results from previous studies [17, 39]. Antisense GapmeR knockdown of MALAT1 RNA remained effective in the MPP^+ treated cells, with a significant loss of MALAT1 occurring after 48 hours of GapmeR treatment (Fig. 6A). After MPP^+ treatment, the expression of the TDP-43 bound mRNA transcripts SLC1A5, CSNK1E, and

HMGB2 were all significantly decreased (Fig. 6B). Knockdown of MALAT1 under these conditions resulted in a significant increase in the levels of each of the TDP-43 bound target mRNA transcripts, indicating that the expression levels of SLC1A5, CSNK1E, and HMGB2 are at least partially dependent upon MALAT1 levels (Fig. 6C). We next repeated the TDP-43 IP-qPCR assays in nuclear extracts of SH-SY5Y neuroblastoma cells treated with MPP^+ and found that TDP-43 binding to MALAT1 was increased, while TDP-43 binding to the 3' UTR of SLC1A5, CSNK1E, and HMGB2 mRNAs was decreased (Fig. 6D).

Finally, we evaluated SH-SY5Y cell viability to determine whether MALAT1 expression contributes to neuroblastoma cell death due to MPP^+ toxicity. Treatment with MPP^+ resulted in death of ~50% of SH-SY5Y cells in culture, consistent with previous reports (Fig. 6E). Targeted depletion of TDP-43 protein and MALAT1 RNA had opposite effects on cell viability. Knockdown of TDP-43 protein decreased cell viability, while knockdown of MALAT1 RNA significantly increased cell viability in MPP^+ treated SH-SY5Y cells (Fig. 6F). Further confirming the direct impact of MALAT1 on cell survival, ectopic overexpression of a full-length MALAT1 RNA transcript was able to overcome the protective effect of MALAT1 GapmeR knockdown after MPP^+ treatment, causing a significant increase in cell death when compared to cells transfected with the empty vector control. Ectopic expression of TDP-43 protein did not have any additional effect on cell death in MPP^+ conditions with MALAT1 knockdown, when compared to cells transfected with the empty vector control (Fig. 6F).

Discussion

Here we investigated the functional effects of modulating MALAT1 RNA levels in HEK293 and SH-SY5Y neuroblastoma cells. We evaluated changes in cell viability and TDP-43

protein function in response to MALAT1 RNA perturbation. Dysregulation of TDP-43 protein function has previously been associated with the progression of multiple neurodegenerative diseases, though the cause of this dysregulation is still unclear. Post-mortem examination of TDP-43 RNA binding in patient brain samples revealed that MALAT1 showed one of the largest increases in RNA binding to TDP-43 compared to healthy control samples [10]. Both TDP-43 and MALAT1 are broadly expressed at high levels in most human cell types and have been proposed to interact with diverse messenger RNAs and splicing factors to affect alternative splicing and gene expression control. Based on these data, we hypothesized that perturbations of MALAT1 could affect cell survival and TDP-43 function in controlling mRNA expression levels. In this study, we found that modulation of MALAT1 levels affects the RNA-binding capacity of TDP-43 for other messenger RNA transcripts in the nucleus, and consequently, affects the expression levels of several mRNA transcripts that are bound with high affinity by TDP-43 in HEK293 and SH-SY5Y cells. MALAT1 and TDP-43 may be linked as part of an RNA-protein interaction network to regulate and maintain appropriate levels of gene expression in human cells.

Studies of TDP-43 binding to RNA genome-wide as well as to specific targets have revealed that TDP-43 is involved in splicing and stability regulation of diverse messenger RNA transcripts. We re-evaluated high-throughput sequencing data from CLIP-seq experiments of TDP-43 binding in three different cell types to identify 1549 commonly bound RNA transcripts in HEK293, SH-SY5Y, and H9 cells. The majority of TDP-43 binding events in this commonly bound transcript set were in the 3’ UTRs, and we selected the SLC1A5, CSNK1E, and HMGB2 for further mechanistic studies. Immunoprecipitation and *in vitro* RNA-protein binding studies validated the identified TDP-43 and mRNA 3’ UTR interactions. We observed that the CLIP

peak height correlated with the calculated TDP-43 *in vitro* binding affinity for the target mRNAs examined (Table 1). TDP-43 binding affinity for these transcripts also correlated with the number of consecutive UG repeats in the mRNA in the bound region, in line with previous work identifying UG repeats as consensus binding sites for TDP-43.

We used a TARDBP mRNA 3' UTR fragment as a positive control for binding studies. As noted in previous reports, the RNA region of TARDBP mRNA where the TDP-43 protein binds for autoregulation control at the 3' UTR of the transcript does not contain numerous sequential UG repeats, despite TDP-43 binding with high affinity and observed peaks in multiple CLIP datasets [33]. Our data support the hypothesis that a lower TDP-43 binding affinity may be desirable for the 3' UTR region in the TARDBP transcript, since multiple molecules may need to be bound to achieve a functional regulatory interaction [33]. Similarly, the MALAT1 RNA transcript region bound to TDP-43 with high affinity *in vitro* and with strong CLIP-seq signal, despite also lacking extensive sequential UG-repeats, additionally supporting a role for MALAT1 in modulating TDP-43 binding at a regulatory level.

Our data show that alterations of MALAT1 RNA levels through either knockdown or overexpression of RNA did not change TDP-43 protein level or cellular localization but resulted in increased binding of TDP-43 to the 3' untranslated region (3' UTR) of messenger RNAs involved in apoptosis. MALAT1 RNA knockdown caused an increase in the expression levels of the three selected target mRNAs, while TDP-43 knockdown had the opposite effect. The role of MALAT1 in disease has been proposed to be context-dependent and could affect gene expression and cell survival through manipulation of complex regulatory pathways [40].

MALAT1 is one of the most abundantly expressed RNA transcripts in mammalian nuclei. RNA binding in healthy cells retains TDP-43 in the nucleus and has been shown to

prevent TDP-43 aggregation [41]. Transient depletion of MALAT1 levels through ASO GapmeR knockdown resulted in increased TDP-43 binding to other mRNA transcripts in the nucleus, without inducing TDP-43 migration to the cytoplasm. In extreme cases of TDP-43 dysfunction, such as RNase treatment or competition with excess UG-containing transcripts, cytoplasmic accumulation of TDP-43 has been observed [42]. The mRNA expression levels of SLC1A5, CSNK1E, and HMGB2 transcripts each increased after MALAT1 depletion, which was caused by TDP-43 binding to and stabilizing the mRNA of these transcripts. Since MALAT1 and TDP-43 levels are coordinated, the MALAT1 expression level may function to affect TDP-43 availability for binding to other transcripts in the nucleus. Modulation of TDP-43 binding to mRNAs could provide a mechanism for MALAT1 to exert a common regulatory effect on different sets of target genes in different cell types.

MALAT1 RNA expression at the appropriate level is required for cell survival. Depletion of MALAT1 caused apoptotic cell death in SH-SY5Y cells, indicating that maintenance of MALAT1 RNA expression promotes neuronal survival. Disruption of MALAT1 lncRNA expression in neuroblastoma cells led to increased binding of TDP-43 to 3' UTRs of mRNA transcripts associated with apoptosis. Treatment with MPP^+ in SH-SY5Y neuroblastoma cells, a model for cell death in neurodegenerative disease, resulted in a strong increase in MALAT1 expression and induced significant cell death. Disruption of MALAT1 expression partially rescued the cell death phenotype of MPP^+ treated neuroblastoma cells, while ectopic overexpression of MALAT1 reversed this protective effect. Localization of TDP-43 to 3' UTR targets of mRNAs was decreased in MPP^+ treated cells, which we hypothesize is due to increased MALAT1 expression and concomitant increased binding of TDP-43 to MALAT1 non-coding RNA transcripts. Since the mechanism of cell death due to MALAT1 overexpression or

protection due to MALAT1 knockdown is still unknown, further research will be required to evaluate the contributions of MALAT1 to disease states including in human neurodegeneration.

In conclusion, these data suggest that MALAT1 RNA expression is required to maintain the correct level of TDP-43 binding to other messenger RNA targets. Our studies confirm the interdependence of TDP-43 and MALAT1 functions and show that dysregulation of non-coding RNA transcripts play a role in cell survival in a cell culture model of neurodegenerative disease. Coordinated regulation and function of MALAT1 and TDP-43 may buffer transient changes in gene expression to maintain appropriate TDP-43 binding to other mRNA transcripts within the nucleus.

Tables and Sequences

Table 2.1 - CLIP-seq peak heights, UG repeat numbers, and Kd of TDP-43 binding targets.

	CLIP-seq peak height SH-SY5Y	CLIP-seq peak height HEK293	CLIP-seq peak height H9	Consecutive UG repeats in RNA region	Kd (nM) from mEMSA
SLC1A5	68	968	306	18	82 ± 11
CSNK1E	16	496	88	14	101 ± 20
HMGB2	23	414	21	6	148 ± 7
BARD1	-	-	-	-	1790 ± 51
TARDBP	15	67	69	2	272 ± 15
MALAT1	310	5190	513	2	127 ± 6

Table 2.2 - RNA sequences for TDP-43 binding sites on mRNAs and MALAT1.

RNA Construct	**Sequence of binding region used for *in vitro* binding assay with TDP-43**
TARDBP$_{[2140-2283]}$ 3'UTR	GAUUGAUGGUGGUGCCGAGGCAUGAAAGGCUAGUAUGAGCGAGA AAAGGAGAGAGCGCGUGCAGAGACUUGGUGGUGCAUAAUGGAUA UUUUUUAACUUGGCGAGAUGUGUCUCUCAAUCCUGUGGCUUUGG UGAGAGAGUGUGCG
MALAT1$_{[6638-6938]}$	AUCCCGCUGCUAUUAGAAUGCAUUGUGAAACGACUGGAGUAUGA UUAAAAGUUGUGUUCCCCAAUGCUUGGAGUAGUGAUUGUUGAAG GAAAAAAUCCAGCUGAGUGAUAAAGGCUGAGUGUUGAGGAAAUU UCUGCAGUUUUAAGCAGUCGUAUUUGUGAUUGAAGCUGAGUACA UUUUGCUGGUGUAUUUUUAGGUAAAAUGCUUUUUGUUCAUUUCU GGUGGUGGGAGGGGACUGAAGCCUUUAGUCUUUUCCAGAUGCAA CCUUAAAAUCAGUGACAAGAAACAUUCCAAACAAGC
SLC1A5$_{[2620-2869]}$ 3'UTR	AACACCAUGCUGGUUAUUUUGGCGGCUGUAGUUGUGGGGGGAUG UGUGUGUGCACGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG UGUGUUCUGUGACCUCCUGUCCCCAUGGUACGUCCCACCCUGUCC CCAGAUCCCCUAUUCCCUCCACAAUAACAGAAACACUCCCAGGGA CUCUGGGGAGAGGCUGAGGACAAAUACCUGCUGUCACUCCAGAGG ACAUUUUUUUUAGCAAUAAAAUUGAG
CSNK1E$_{[2054-2264]}$ 3'UTR	GUGGGCUUUUCCAUUGUCCCCCUGGCCUCCAGGCUCCUCCUCUGC CUCUCCAUGGAGUGGGUGGGGAGGUGGUGGGGGCCGGCGUCCCCU GCGUGUGUGUGUGUGUGUGUGUGUGUGUGGAUGUAUUGACCU GUGUUUCCCAAGACAGCAGGUGCCACGGCCCGCCCCGCCUGCCAG CCCGAAUUCCCGUUCUCCUGUGUCUACUAACAAGGACGGAUCCAG CU
HMGB2$_{[739-1022]}$ 3'UTR	UGGCUAUCCUUUAAUGAUGCGUGUGGAAUGUGUGUGUGUGCUCA GGCAAUUAUUUUGCUAAGAAUGUGAAUUCAAGUGCAGCUCAAUA CUAGCUUCAGUAUAAAAACUGUACAGAUUUUUGUAUAGCUGAUA AGAUUCUCUGUAGAGAAAAUACUUUUAAAAAAUGCAGGUUGUAG CUUUUUGAUGGGCUACUCAUACAGUUAGAUUUUACAGCUUCUGA UGUUGAAUGUUCCUAAAUAUUUAAUGGUUUUUUUAAUUUCUUGU GUAUGGUAGCACAGCAAAGCUU
BARD1$_{[86-310]}$ Negative Control	GCGAGGAGCCUUUCAUCCGAAGGCGGGACGAUGCCGGAUAAUCGG CAGCCGAGGAACCGGCAGCCGAGGAUCCGCUCCGGGAACGAGCCU CGUUCCGCGCCCGCCAUGGAACCGGAUGGUCGCGGUGCCUGGGCC CACAGUCGCGCCGCGCUCGACCGCCUGGAGAAGCUGCUGCGCUGC UCGCGUUGUACUAACAUUCUGAGAGAGCC

Table 2.3 - Oligonucleotide primers used for quantitative RT-PCR analysis.

Primer Name	Sequence (5' to 3')
SLC1A5 FWD	CTCTTCACCCGCAAAAACCC
SLC1A5 REV	GCTTGGCCACGCCATTATTC
CSNK1E FWD	GAATTCCCGTTCTCCTGTGTCTA
CSNK1E REV	AAAACCAGGAATGGAAGATGGAG
HMGB2 FWD	TGCTCTGAACATCGCCCAAA
HMGB2 REV	GCTTCACTTTTGCCCTTGGC
TARDBP FWD	TCAGGGCCTTTGCCTTTGTT
TARDBP REV	TGCTTAGGTTCGCATTGGAT
MALAT1 FWD	GAAGGAAGGAGCGCTAACGA
MALAT1 REV	TACCAACCACTCGCTTTCCC
GAPDH FWD	GGGCTCTCCAGAACATCATCC
GAPDH REV	GTCCACCACTGACACGTTGG
TUBULIN FWD	CCAGACAACTTTGTATTTGGTCAGT
TUBULIN REV	CGTACCACATCCAGGACAGAAT
BARD1 FWD	GCCAAAGCTGTTTGATGGAT
BARD1 REV	CGAACCCTCTCTGGGTGATA

Methods

Purification of TDP-43 RNA recognition motif domains

The RNA recognition motifs (RRMs) (Q101-Q269) of TDP-43 were purified from a construct with an N terminus 6x-His-Gb1 tag and a TEV cleavage site. This construct was transformed into BL21 *E. coli* and expression was induced with 2% (w/v) L-arabinose at OD600 ~0.7 with in LB medium with ampicillin at 16°C overnight. The fusion protein was purified with a HisTrap FF 1 mL column (Cytiva 17-5319-01) with Buffer A (20 mM HEPES, pH 7.5, 1 M NaCl, 10% glycerol (w/v), 30 mM imidazole and 5 mM 2-mercaptoethanol) and Buffer B (same as Buffer A, but containing 300 mM imidazole) on an AKTA pure FPLC system at 4 °C (Cytiva). The eluted protein was dialyzed overnight at 4 °C in Buffer A, then cleaved with TEV protease [43] for 4 hours at 30 °C. The protein was reapplied to the HisTrap FF column and the flow through containing untagged TDP-43 was collected. The purified protein was concentrated

with a 10kDa MWCO centrifugal concentrator (Thermo 88527) and buffer exchanged into 50mM HEPES pH 7.5, 300mM NaCl, 10% glycerol, and 5mM beta-mercaptoethanol before being flash frozen in liquid nitrogen and stored at –80 °C.

Molecular cloning

Primers for the genes of interest were designed using NIH's NCBI Primer-BLAST software and were synthesized by Integrated DNA Technologies. Genes of interest were amplified by polymerase chain reaction (PCR) using cDNA from K562 cells (ATCC CCL-243) and Phusion polymerase (NEB M0530S). The PCR products were run on a 2% agarose gel stained with SYBR Safe (Thermo S33102) to confirm PCR efficacy. The PCR product was cleaned up using Qiagen QIAquick PCR Purification Kit (28104). The genes of interest were ligated into pUC57 plasmids and transformed into DH5α competent cells (Thermo EC0112). Bacterial colonies that demonstrated growth were sampled and then grown overnight in LB growth media. pUC57 plasmids were isolated from the transformation product using the Qiagen QIAprep Spin Miniprep Kit (27106X4). Plasmid concentrations were measured with the Eppendorf BioSpectrometer basic (6135) and fully sequenced for the genes of interest through Genewiz-Azenta (La Jolla, CA, USA).

CLIP-seq data analysis

Overlap in TDP-43 binding on RNA transcripts from different cell types was identified from the datasets SRR4044755 from H9 cells [21], ERR039855 from SH-SY5Y neuroblastoma cells [10], and ERR9192743 from 293HEK embryonic kidney cells [22]. The datasets with UMIs were processed with UMI Tools [44]. All datasets were trimmed with Cutadapt [45] to remove

adapters, low quality bases, and short sequences (<20 bp). Datasets with barcodes were filtered only for reads with the correct barcode. Reads were aligned to the human genome (hg38) with STAR [46], and the alignment files were indexed with samtools [47]. TPMs were calculated with TPMCalculator [48], and the top 5000 genes in each dataset were used in further analysis. Reads in the three datasets were overlapped with bedtools [49]. The overlapped reads were analyzed with HOMER findPeaks to find binding sites [50]. Gene binding site enrichment was generated with HOMER annotatePeaks.pl. Gene ontology analysis of the commonly bound transcripts was performed with ShinyGO [51]. Graphic representations of CLIP peaks for each gene bound by TDP-43 were created from the SH-SY5Y iCLIP data set ERR039855 using the Integrative Genomics Viewer.

Multiplexed electrophoretic mobility shift assays (mEMSA)

Multiplexed EMSA experiments were performed as previously described [32]. The mEMSA binding studies were performed with purified TDP-43 protein RRMs and fluorescently labeled *in vitro* transcribed RNA fragments. RNA transcripts were generated by *in vitro* transcription using NEB HiScribe T7 High Yield RNA synthesis kit (E2040S), with the addition of fluorophore labeled UTPs: Cyanine 3-uridine-5'-triphosphate (enhanced) from Enzo Life Sciences, Inc. (catalog #ENZ-42505), fluorescein-12-uridine-5'-triphosphate from Enzo Life Sciences, Inc. (catalog #ENZ-42834), and cyanine 5-UTP from Apexbio Technology LLC (purchased through Fisher Scientific catalog # 50-199-8343). The RNA sequences used for binding studies are listed in Table 2.

Sets of labeled RNA targets with unique fluorophores were then mixed in equimolar (2 nM each) amounts and heat refolded. Increasing amounts of TDP-43 were incubated with the

RNA for 30 minutes at room temperature and then run on a 5.5% TBE polyacrylamide gel. Gels were imaged on a Typhoon FLA 9500 with pixel size of 50um for each fluorophore. The lasers and filter combinations were 473nm/BPB1, 532nm/BPG1, and 635 nm/LPR for fluorescein Cy3, and Cy5, respectively. Gel images were qualified with ImageStudioLite. The dissociation constant, Kd, was calculated by fitting to Equation (1) using the nonlinear least squares method in R:

$$(1)\ Fraction\ bound\ = \frac{[P]^n}{[P]^n + K_d}$$

Cell culture and transfections

HEK293 cells (ATCC CRL-1573) were grown in T-75 flasks in Eagle's Minimum Essential Medium (EMEM). SH-SY5Y neuroblastoma cells (ATCC CRL-2266) were grown in T-75 flasks in a 1:1 ratio of EMEM with Gibco Ham's F12 Media. All media were supplemented with 10% fetal bovine serum (FBS) and 1% L-glutamine. Cells were maintained at 37°C in incubation with 5% CO_2 atmosphere and routinely tested for mycoplasma contamination. Transfections with locked nucleic acid (LNA) GapmeRs were used to achieve transient knockdown of the target RNA and avoid affecting transcription or enhancer function at the DNA locus. GapmeRs with antisense sequence specific to MALAT1 (LG0000003-DDA) and a non-targeting control GapmeR (LG00000002-DDA) were purchased from Qiagen. RNA knockdowns were performed as per manufacturer protocols, using OptiMeM 1 Reduced Serum Medium. HEK293 cells were transfected with LNA GapmeRs using Lipofectamine 2000 for 24 hours, while SH-SY5Y cells were transfected using Lipofectamine 3000 for 48 hours. Cells were collected, centrifuged, and assessed for cell viability. Knockdown efficiency was verified

through RT-qPCR. Overexpression plasmids were transfected with Lipofectamine 3000 and P3000 reagent for 48 hours.

Cell lysis and RNA immunoprecipitation

HEK-293 cells were grown in T-75 flasks and transfected with MALAT1 and NC GapmeR respectively for 24 hours. Cells were collected and resuspended in 1mL of HLB buffer containing 10mM Tris (pH 7.5), 10mM NaCl, 3mM $MgCl_2$, 0.3% (v/v) NP-40 and 10% (v/v) glycerol with 1× EDTA- free Protease Inhibitor Cocktail (PIC). After 10 mins of incubation at 4 °C, cells were briefly vortexed and centrifuged at 800 x *g* for 8 min. The supernatant was collected, and the pellet was washed three times with 100 µL 1× PLB and centrifuged at 200 x *g* for 2 min. The nuclear pellet was resuspended in 1× NLB buffer (20 mM Tris pH 7.5, 150 mM KCl, 3 mM $MgCl_2$, 0.3% (v/v) NP-40 and 10% (v/v) glycerol) with 1× PIC. Nuclei were sonicated three times at 20% power for 15 sec in an ice bath with 2 min cooling between each sonication. Resuspended nuclei and collected cytoplasmic supernatant were centrifuged at 18,000 x *g* for 15 min and final supernatants were pooled together. A sample of 20 µL of lysate was saved for RIP input analysis. RNA immunoprecipitation was performed as described previously [36]. The following buffers were prepared for RNA-IP: 5X NT-2 buffer (250 mM Tris-HCl pH 7.4, 750 mM NaCl, 5 mM $MgCl_2$, 0.25 % NP-40), and NET-2 Buffer (1X NT-2 buffer supplemented with 20 mM EDTA pH 8.0, 1 mM DTT, 200 units/ml RNase OUT). For each capture, 75µL of Dynabeads Protein-G were incubated with 50 µg of antibodies corresponding to TDP-43 (10782-2-AP) and IgG (30000-0-AP) for 1.5 hours at RT. Beads were washed with 1×-NT-2 buffer and the pooled and lysed supernatant were added to each sample with NET-2 buffer and rotated overnight at 4°C. Samples were then washed 6 times with 1× NT-

2 buffer and split into equal halves for protein and RNA extraction. Proteins were extracted from the beads through elution with 2X SDS buffer and ran on a 12% PAGE gel to assess immunoprecipitation efficiency. RNA was extracted and purified through ethanol precipitation onto silane-treated magnetic beads. Purified RNA was resuspended in 10 µL UP H_2O and was processed for RT-qPCR.

Fluorescence Imaging

Immunofluorescence was conducted essentially as previously described, with the following modifications [52]. SH-SY5Y cells were seeded on a glass-bottomed tissue culture plate and incubated overnight at 37°C. Cells were transfected accordingly with NC or MALAT1 GapmeR for 48 hours. Post transfection, cells were washed 3 times in 1X PBS and fixed with 4% PFA in 1X PBS for 20 minutes at room temperature. Cells were then washed twice, followed by permeabilization in 0.2% Triton X-100 in PBS for 10 minutes, followed by 3 washes in PBS. Cells were blocked in 1% BSA and Native Goat Serum in 1X PBST for 30minutes, followed by 3 washes with PBST. Cells were incubated with the TDP-43 primary antibody diluted in PBST overnight, with dilutions performed as per manufacturer's recommendation. Post incubation, cells were washed three times and incubated in diluted secondary antibody in TBST for 1 hour at room temperature in the dark, followed by three washes with PBS for 5 minutes. For visualization, cells were incubated in 1 µg/mL DAPI DNA stain (4083S, Cell Signaling Technology) or 0.5 µg/ml Hoechst 33342 Reagent (R37165, ThermoFisher Scientific, USA) for 1 min, rinsed with PBS, and imaged using an Eclipse Ti Inverted Microscope (Nikon, Japan) with Nikon's NIS-Elements imaging software.

For nuclear morphology analysis, cell samples were plated on to glass bottom dishes and transfected as mentioned previously with NC or MALAT1 GapmeR for 48 hours. Cells were fixed using 4% paraformaldehyde (Electron Microscopy Sciences, USA) at room temperature for 10 minutes, followed by a minimum of three washes with PBS. Cells were permeabilized with 1% with Triton X-100 (Sigma-Aldrich, USA) for 10 minutes and blocked in 2% Bovine Serum Albumin (BSA) (Thermo Scientific™, USA) for 1 hour. Cells were incubated with 4',6-diamidino-2-phenylindole (DAPI) (Thermo Scientific™, USA) for 5 minutes and Alexa Fluor® 555 Phalloidin (Invitrogen™, USA) for 20 minutes in the dark, to stain the nucleus and F-actin, respectively. Imaging of the samples was performed using an Echo Revolution microscope, utilizing 40x plan fluorite phase objective with NA .75. Nuclear morphology analysis and quantification were performed using Nuclear Morphology Analysis (NMA) in ImageJ [19] to determine the change in overall nuclear area (in pixels) and classification of nuclei across the two conditions.

Real-time quantitative PCR (RT-qPCR) assays

Total RNA was isolated from 1 x 10^6 cells using the RNeasy Mini Kit and resuspended in 30 µL of Invitrogen UltraPure H_2O. cDNA was synthesized using 1 µg of RNA and the ProtoScript II Reverse Transcriptase, using the protocol provided by the manufacturer. qPCR was performed on a QuantStudio 3 using appropriate primers and a housekeeping gene target for control. The knockdown efficiency was calculated using the ΔC_t method and was measured in technical triplicates. RIP Fold Change was calculated by determining the % of RNA compared to input sample as previously described [36]. All primer sequences used for qPCR are listed in Table 3.

Cell viability assays

Cell viability was measured through CCK8 (Abcam) assays according to the manufacturer's protocol. SH-SY5Y cells were seeded onto a 12-well plate and transfected with antisense oligonucleotide (ASO) locked nucleic acid (LNA) GapmeRs and 1-methyl-4-phenylpyridinium (MPP^+) for 48 hours. After transfection, each well was incubated with 6 μL of CCK8 reagent and allowed to incubate at 37°C for 2 hours. 96 well plates were shaken for 30 seconds, followed by imaging at 460 nm using a Synergy HTX multimode microplate reader.

Western blotting

For protein isolation and quantification, 3×10^5 SH-SY5Y cells were seeded per well on 12 well tissue culture treated plates and cells were transfected with Lipofectamine 3000 in combination with GapmeR or overexpression plasmids for the indicated amount of time. Cells were lysed in 1X SDS buffer post transfection and boiled to 95°C to complete lysis before SDS-PAGE separation. A 12% Bis-acrylamide Tris Buffer SDS-PAGE Gel was run at 130 volts for 90 minutes to separate proteins, which were subsequently transferred onto a nitrocellulose membrane for 50 minutes at 12 volts. Post transfer, the membrane was blocked in 5% nonfat dry milk diluted in 10 mL 1X TBS for 1 hour at room temperature. Three washes with 1X TBS were performed for 5 minutes at room temperature. The membrane was incubated overnight at 4 °C with primary antibody dilution in 10mL 1X TBST. The membrane was washed 3 times with 1X TBST and subsequently incubated in the indicated secondary antibody dilution in 10 mL 1XTBST for 1 hour at room temperature. After three final washes, the membrane was imaged at 700 nm and 800 nm on a Li-Cor Odyssey FC imager with a two-minute acquisition time. Image

analysis and quantification were performed with Empiria Studio using automatic background correction. All protein quantifications were normalized to an internal loading control in the same lane. In these experiments, the primary antibodies used were: TDP-43 Anti Rabbit (10782-2-AP) at a 1:3000 dilution, and GAPDH Anti Mouse (60004-1-Ig). The following IR-dye conjugated secondary antibodies were used: Goat-anti mouse conjugated to 800nm IR Dye and Goat anti-rabbit conjugated to 700nm IR Dye. Antibodies were validated before use by staining serial dilutions of total cell lysate.

For Western blot analysis of protein elutions in RNA Immunoprecipitation experiments, the above protocol was slightly modified. A VeriBlot secondary antibody for IP detection from Abcam (ab131366) was used, and visualization was performed using the Clarity ECL Western substrate. Imaging was performed on the Li-Cor Odyssey FC imager in the chemiluminescence channel.

References

1. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun. 2006 Dec 22;351(3):602-11. doi: 10.1016/j.bbrc.2006.10.093. Epub 2006 Oct 30. PMID: 17084815.

2. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006 Oct 6;314(5796):130-3. doi: 10.1126/science.1134108. PMID: 17023659.

3. Hasegawa M, Arai T, Akiyama H, Nonaka T, Mori H, Hashimoto T, Yamazaki M, Oyanagi K. TDP-43 is deposited in the Guam parkinsonism-dementia complex brains. Brain. 2007 May;130(Pt 5):1386-94. doi: 10.1093/brain/awm065. Epub 2007 Apr 17. PMID: 17439983.

4. Nakashima-Yasuda H, Uryu K, Robinson J, Xie SX, Hurtig H, Duda JE, Arnold SE, Siderowf A, Grossman M, Leverenz JB, Woltjer R, Lopez OL, Hamilton R, Tsuang DW, Galasko D, Masliah E, Kaye J, Clark CM, Montine TJ, Lee VM, Trojanowski JQ. Co-morbidity of TDP-43

proteinopathy in Lewy body related diseases. Acta Neuropathol. 2007 Sep;114(3):221-9. doi: 10.1007/s00401-007-0261-2. Epub 2007 Jul 25. PMID: 17653732.

5. Nelson, P.T., Dickson, D.W., Trojanowski, J.Q., Jack, C.R., Boyle, P.A., Arfanakis, K., Rademakers, R., Alafuzoff, I., Attems, J., Brayne, C. and Coyle-Gilchrist, I.T., 2019. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. Brain, 142(6), pp.1503-1527.

6. Buratti, E., and Baralle, F. E. (2001) Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J. Biol. Chem.* **276**, 36337-36343

7. Volkening, K., Leystra-Lantz, C., Yang, W., Jaffee, H., and Strong, M. J. (2009) Tar DNA binding protein of 43 kDa (TDP-43), 14-3-3 proteins and copper/zinc superoxide dismutase (SOD1) interact to modulate NFL mRNA stability. Implications for altered RNA processing in amyotrophic lateral sclerosis (ALS). *Brain Res.* **1305**, 168–182

8. Mann, J.R., Gleixner, A.M., Mauna, J.C., Gomes, E., DeChellis-Marks, M.R., Needham, P.G., Copley, K.E., Hurtle, B., Portz, B., Pyles, N.J. and Guo, L., 2019. RNA binding antagonizes neurotoxic phase transitions of TDP-43. Neuron, 102(2), pp.321-338.

9. Loganathan, S., Lehmkuhl, E. M., Eck, R. J., and Zarnescu, D. C. (2020) To Be or Not To Be…Toxic—Is RNA Association With TDP-43 Complexes Deleterious or Protective in Neurodegeneration? *Front. Mol. Biosci.* **8,** 154

10. Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Župunski, V. and Patani, R., 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nature neuroscience, 14(4), pp.452-458.

11. Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J. B., and Chess, A. (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics*. **8**, 39

12. Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A. and Blencowe, B.J., 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell, 39(6), pp.925-938.

13. Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E. and Thomas, M., 2003. MALAT-1, a novel noncoding RNA, and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene, 22(39), pp.8031-8041.

14. Bernard, D., Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M.Q., Sedel, F., Jourdren, L., Coulpier, F. and Triller, A., 2010. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. The EMBO journal, 29(18), pp.3082-3093.

15. Guo, F., Jiao, F., Song, Z., Li, S., Liu, B., Yang, H., Zhou, Q. and Li, Z., 2015. Regulation of MALAT1 expression by TDP43 controls the migration and invasion of non-small cell lung cancer cells in vitro. Biochemical and biophysical research communications, 465(2), pp.293-298.

16. West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y. and Kingston, R.E., 2014. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Molecular cell, 55(5), pp.791-802.

17. Liu, W., Zhang, Q., Zhang, J., Pan, W., Zhao, J., and Xu, Y. (2017) Long non-coding RNA MALAT1 contributes to cell apoptosis by sponging miR-124 in Parkinson disease. *Cell Biosci*. **7**, 19

18. Nguyen, T.M., Kabotyanski, E.B., Reineke, L.C., Shao, J., Xiong, F., Lee, J.H., Dubrulle, J., Johnson, H., Stossi, F., Tsoi, P.S. and Choi, K.J., 2020. The SINEB1 element in the long non-coding RNA Malat1 is necessary for TDP-43 proteostasis. Nucleic acids research, 48(5), pp.2621-2642.

19. Filippi-Chiela, E. C., Oliveira, M. M., Jurkovski, B., Callegari-Jacques, S. M., da Silva, V. D., and Lenz, G. (2012) Nuclear morphometric analysis (NMA): screening of senescence, apoptosis, and nuclear irregularities. *PLoS One*. **7**, e42522

20. Suk, T. R., and Rosseaux, M. W. C. (2020) The role of TDP-43 mislocalization in amyotrophic lateral sclerosis. *Mol. Neurodegener*. **15**, 45

21. Wu, H., Yin, Q.F., Luo, Z., Yao, R.W., Zheng, C.C., Zhang, J., Xiang, J.F., Yang, L. and Chen, L.L., 2016. Unusual processing generates SPA LncRNAs that sequester multiple RNA binding proteins. Molecular cell, 64(3), pp.534-548.

22. Hallegger, M., Chakrabarti, A.M., Lee, F.C., Lee, B.L., Amalietti, A.G., Odeh, H.M., Copley, K.E., Rubien, J.D., Portz, B., Kuret, K. and Huppertz, I., 2021. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell, 184(18), pp.4680-4696.

23. Fiesel, F.C., Voigt, A., Weber, S.S., Van den Haute, C., Waldenmaier, A., Görner, K., Walter, M., Anderson, M.L., Kern, J.V., Rasse, T.M. and Schmidt, T., 2010. Knockdown of transactive response DNA-binding protein (TDP-43) downregulates histone deacetylase 6. The EMBO journal, 29(1), pp.209-221.

24. Yu, Z., Fan, D., Gui, B., Shi, L., Xuan, C., Shan, L., Wang, Q., Shang, Y. and Wang, Y., 2012. Neurodegeneration-associated TDP-43 interacts with fragile X mental retardation protein (FMRP)/Staufen (STAU1) and regulates SIRT1 expression in neuronal cells. Journal of Biological Chemistry, 287(27), pp.22560-22572.

25. Liu, B.W., Wang, X.Y., Cao, J.L., Chen, L.L., Wang, Y.L., Zhao, B.Q., Zhou, J. and Shen, Z.F., 2022. TDP-43 upregulates lipid metabolism modulator ABHD2 to suppress apoptosis in hepatocellular carcinoma. Communications Biology, 5(1), p.816.

26. Abraham, A. B., Bronstein, R., Reddy, A. S., Maletic-Savatic, M., Aguirre, A., and Tsirka, S. E. (2013) Aberrant neural stem cell proliferation and increased adult neurogenesis in mice lacking chromatin protein HMGB2. *PLoS One.* **8**, e84838

27. Lee, N. Y., Kim, Y., Ryu, H., and Kang, Y. S. (2017) The alteration of serine transporter activity in a cell line model of amyotrophic lateral sclerosis (ALS). *Biochem. Biophys. Res. Commun.* **483**, 135-141

28. Adler, P., Mayne, J., Walker, K., Ning, Z., and Figeys, D. (2019) Therapeutic Targeting of Casein Kinase 1δ/ε in an Alzheimer's Disease Mouse Model. *J. Proteome Res.* **18**, 3383-3393

29. Perez, D. I., Gil, C., and Martinez, A. (2011) Protein kinases CK1 and CK2 as new targets for neurodegenerative diseases. *Med. Res. Rev.* **31**, 924-54

30. Krach, F., Batra, R., Wheeler, E.C., Vu, A.Q., Wang, R., Hutt, K., Rabin, S.J., Baughn, M.W., Libby, R.T., Diaz-Garcia, S. and Stauffer, J., 2018. Transcriptome–pathology correlation identifies interplay between TDP-43 and the expression of its kinase CK1E in sporadic ALS. Acta neuropathologica, 136, pp.405-423.

31. Lukavsky, P.J., Daujotyte, D., Tollervey, J.R., Ule, J., Stuani, C., Buratti, E., Baralle, F.E., Damberger, F.F. and Allain, F.H., 2013. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. Nature structural & molecular biology, 20(12), pp.1443-1449.

32. Button, A. C., Hall, S. D., Ashley, E. L., and McHugh, C. A. (2024) Dissection of protein and RNA regions required for SPEN binding to XIST A-repeat RNA. *RNA.* **30**, 240-255

33. Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E. and Baralle, F.E., 2011. TDP-43 regulates its mRNA levels through a negative feedback loop. The EMBO journal, 30(2), pp.277-288.

34. Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R., Huelga, S.C., Moran, J., Liang, T.Y., Ling, S.C., Sun, E., Wancewicz, E., Mazur, C. and Kordasiewicz, H., 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nature neuroscience, 14(4), pp.459-468.

35. Sephton, C.F., Cenik, C., Kucukural, A., Dammer, E.B., Cenik, B., Han, Y., Dewey, C.M., Roth, F.P., Herz, J., Peng, J. and Moore, M.J., 2011. Identification of Neuronal RNA Targets of TDP-43-containing Ribonucleoprotein Complexes*♦. Journal of biological chemistry, 286(2), pp.1204-1215.

36. Gagliardi, M. and Matarazzo, M.R., 2016. Rip: Rna Immunoprecipitation. Polycomb Group Proteins: Methods and Protocols, pp.73-86.

37. Kalivendi, S. V., Kotamraju, S., Cunningham, S., Shang, T., Hillard, C. J., and Kalyanaraman, B. (2003) 1-Methyl-4-phenylpyridinium (MPP+)-induced apoptosis and mitochondrial oxidant generation: role of transferrin-receptor-dependent iron and hydrogen peroxide. *Biochem. J.* **371**, 151-164

38. Cassarino, D. S., Parks, J. K., Parker, W. D. Jr., and Bennett, J. P. Jr. (1999) The parkinsonian neurotoxin MPP+ opens the mitochondrial permeability transition pore and releases cytochrome c in isolated mitochondria via an oxidative mechanism. *Biochim. Biophys. Acta*. **1453**, 49-62

39. Lu, Y., Gong, Z., Jin, X., Zhao, P., Zhang, Y., and Wang, Z. (2020) LncRNA MALAT1 targeting miR-124-3p regulates DAPK1 expression contributes to cell apoptosis in Parkinson's Disease. *J. Cell. Biochem*. **121**, 4838-4848

40. Arun, G., Aggarwal, D., and Spector, D. L. (2020) *MALAT1* Long Non-Coding RNA: Functional Implications. *Noncoding RNA.* **6**, 22

41. Rengifo-Gonzalez, J.C., El Hage, K., Clément, M.J., Steiner, E., Joshi, V., Craveur, P., Durand, D., Pastré, D. and Bouhss, A., 2021. The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. Elife, 10, p.e67605.

42. Duan, L., Zaepfel, B.L., Aksenova, V., Dasso, M., Rothstein, J.D., Kalab, P. and Hayes, L.R., 2022. Nuclear RNA binding regulates TDP-43 nuclear localization and passive nuclear export. Cell reports, 40(3).

43. Raran-Kurussi, S., Cherry, S., Zhang, D., and Waugh, D. S. (2017) Removal of affinity tags with TEV protease. *Methods Mol. Biol.* **1586**, 221-230.

44. Smith, T., Heger, A., and Sudbery, I. (2017) UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491-499

45. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal,* **17**, 10-12

46. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), pp.15-21.

47. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. and Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience, 10(2), p.giab008.

48. Alvarez, R. V., Pongor, L. S., Mariño-Ramírez, L., and Landsman, D. (2019) TPMCalculator: one-step software to quantify mRNA abundance of genomic features. *Bioinformatics*. **35**, 1960–1962

49. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**: 841–842

50. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell, 38(4), pp.576-589.

51. Ge, S. X., Jung, D., and Yao, R. (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. **36**, 2628-2629.

52. [preprint] Su, T., Trang, N., Zhu, J., Kong, L., Cheung, D., Chou, V., Ellis, L., Huang, C., Camden, N. and McHugh, C.A., 2023. GRAS1 non-coding RNA protects against DNA damage and cell death by binding and stabilizing NKAP. bioRxiv, pp.2023-06.

Acknowledgements

We appreciate the continued support of our sponsors and the Department of Chemistry and Biochemistry at UCSD.

This work was supported by NIH grant R00 GM120494 to C.A.M., Molecular Biophysics Training Grant support T32 GM008326 to A.C.B., the San Diego Fellowship to A.B., UCSD Chemistry and Biochemistry faculty startup funds to C.A.M, and the UCSD Dean of Physical Sciences Summer Undergraduate Research Award to R.J.

Chapter 2, in full, has been submitted for publication of the material as it may appear in the Journal of Biological Chemistry. This material includes author contributions from Aileen C. Button, Jonathan Zhu, Lauren Ellis, Ellen Lavorando, Ethan L. Ashley, Raul Johnson, Einollah Sarikhani, Zeinab Jahed, and Colleen A. McHugh. The dissertation author was the primary researcher and author of this material.

Figures

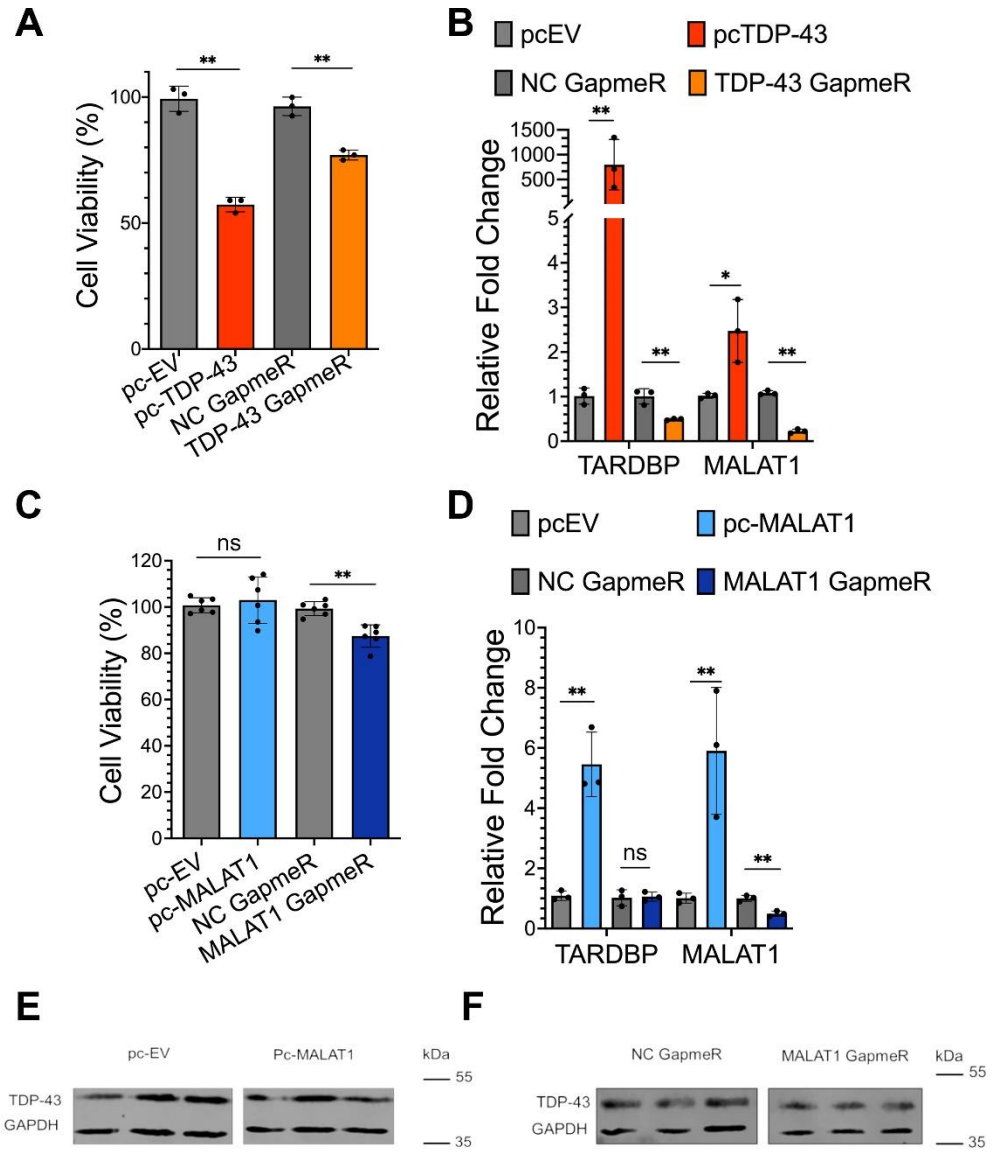


Figure 2.1 - Disruption of MALAT1 or TDP-43 induces cell death and perturbs RNA transcript levels of the other binding partner in SH-SY5Y neuroblastoma cells.

(A) Cell viability after overexpression or knockdown of TDP-43. (B) Cell viability after overexpression or knockdown of MALAT1 non-coding RNA. (C) Gene expression changes in TDP-43 and MALAT1 RNA levels, after TDP-43 overexpression or knockdown. (D) Gene expression changes in TDP-43 and MALAT1 RNA levels, after MALAT1 overexpression or knockdown. (E) Western blot for protein expression changes in TDP-43 after MALAT1 overexpression (F) Western blot for protein expression changes in TDP-43 after MALAT1 knockdown. All experiments were performed in SH-SY5Y cell lines. N=3-6 for cell viability experiments, N=3 for gene expression and western blot experiments. T tests are conducted with two tailed unpaired equal variance conditions. * = $p<0.05$, ** = $p<0.01$. All data are plotted with standard deviation.

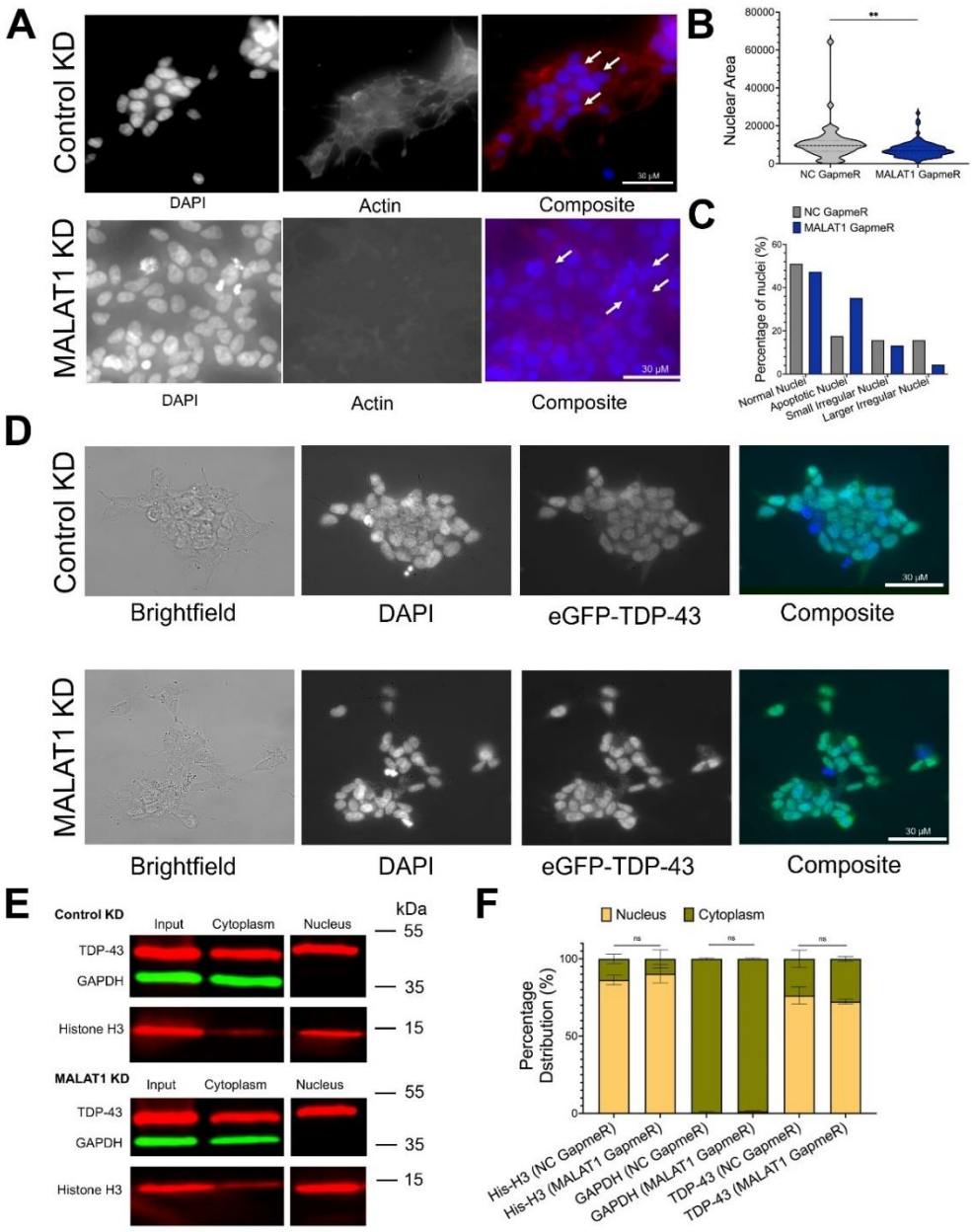


Figure 2.2 - MALAT1 KD induces cell death in SH-SY5Y cells but does not lead to cytoplasmic localization of TDP-43.

(A) Immunofluorescence imaging of actin with DAPI nuclear staining, in SH-SY5Y cells after 48 hours of transfection with negative control (NC) GapmeR or MALAT1 GapmeR. Scale bar, 20 µM. N=51 for NC GapmeR, N=91 for MALAT1 GapmeR. Arrows indicate examples of normal nuclei in control knockdown and aberrant nuclei in MALAT1 knockdown treated cells. (B) Quantification of nuclear area in NC GapmeR or MALAT1 GapmeR treated cells. (C) Classification of apoptotic and aberrant nuclei in NC GapmeR or MALAT1 GapmeR treated cells. (D) Immunofluorescence imaging of TDP-43 with DAPI nuclear staining in SH-SY5Y cells after 48 hours of transfection with NC or MALAT1 GapmeR. Scale bars = 30 µM (E) Western blot analysis of TDP-43 protein level in cytoplasmic and nuclear fractions purified from HEK293 cells transfected for 24 hours with NC or MALAT1 GapmeR. N = 3 biological replicates. (F) Quantification of TDP-43 distribution in nuclear and cytoplasmic fractions from Western blot data. T tests are conducted with two tailed unpaired equal variance conditions. * = $p<0.05$, ** = $p<0.01$. All data are plotted with standard deviation.

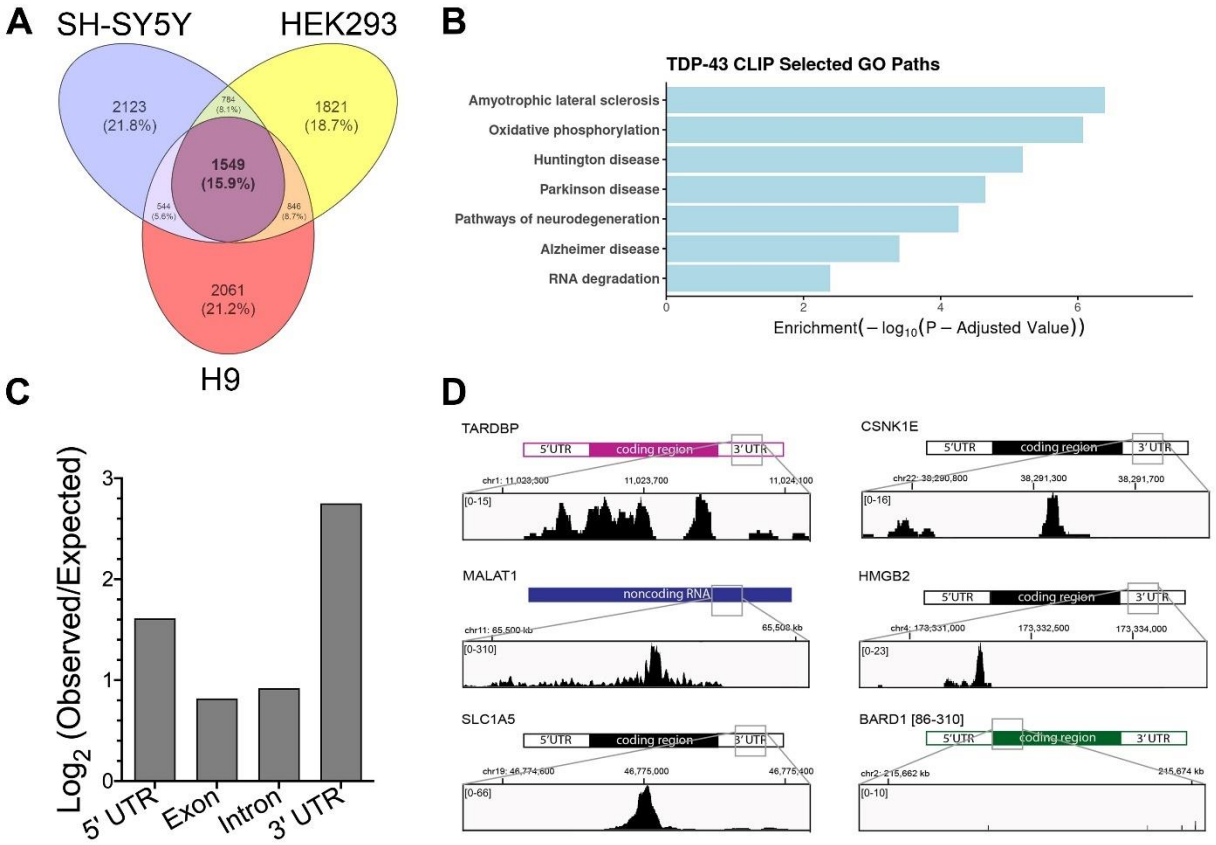


Figure 2.3 - TDP-43 binds the 3' UTR of mRNA targets in multiple cell types.

(A) Overlap of top 5000 transcripts bound by TDP-43 from CLIP-seq datasets from SH-SY5Y neuroblastoma cells (ERR039855), H9 human embryonic stem cells (SRR4044755), and HEK293 cells (ERR9192743). (B) Gene ontology analysis of the 1549 commonly bound genes among the three CLIP-seq datasets. (C) Log fold enrichment of TDP-43 binding regions in the commonly bound transcript set. (D) Integrated Genomics Viewer CLIP-seq read profiles of TDP-43 binding to HMGB2, SLC1A5, CSNK1E, the negative control BARD1 transcript, and MALAT1 RNA regions. Reads were visualized from the SH-SY5Y TDP-43 CLIP-Seq dataset ERR039855.

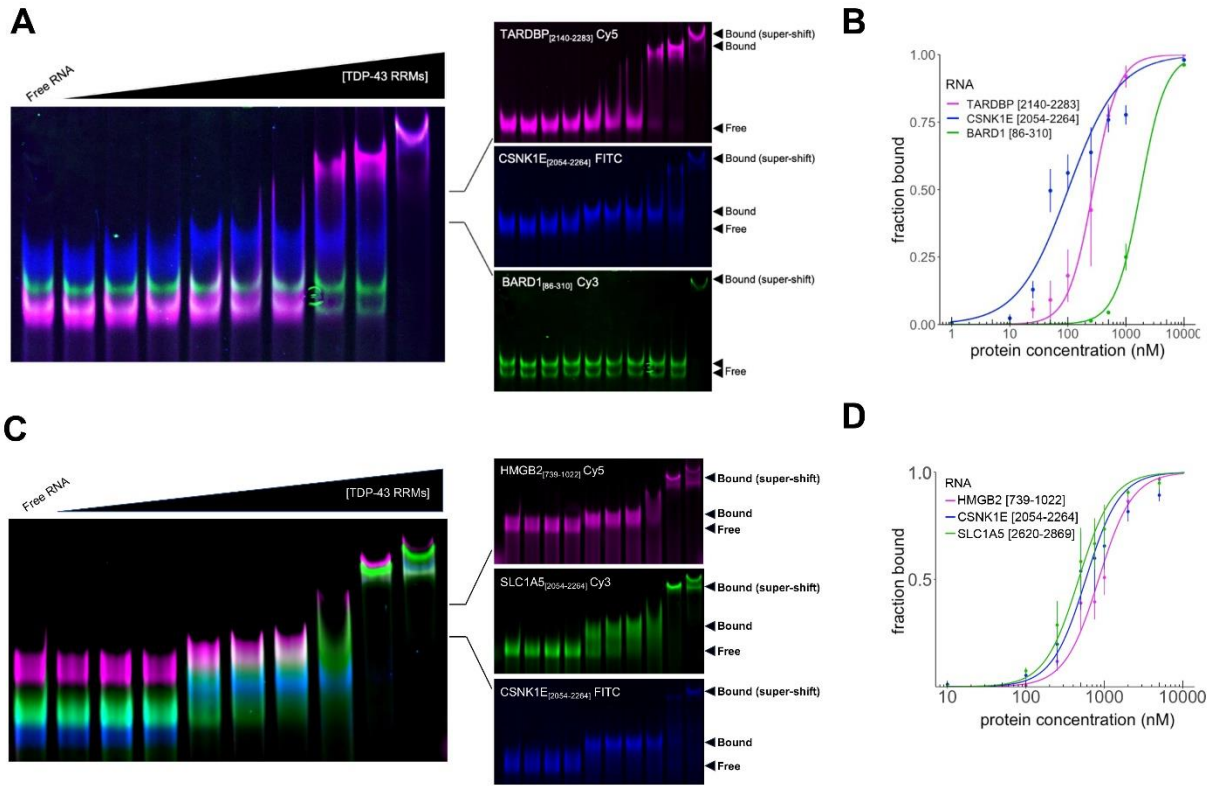


Figure 2.4 - TDP-43 binds the 3' UTR of messenger RNAs with high affinity in vitro.

(A) Representative image from the multiplexed electrophoretic mobility shift assay (mEMSA) for TDP-43 binding messenger RNA 3' UTR fragment of CSNK1E, positive control TARDP mRNA region, and negative control BARD1 mRNA region. (B) Hill's plot for TDP-43 binding with CSNK1E, TARDBP, and BARD1 RNA transcripts. (C) Representative image from the mEMSA for TDP-43 binding to messenger RNA 3' UTR fragments from HMGB2, SLC1A5, and CSNK1E. (D) Hill's plot for TDP-43 binding with messenger RNA 3'UTR transcripts. Data are plotted with standard deviation. N=3 to 6 replicates for each sample pair.

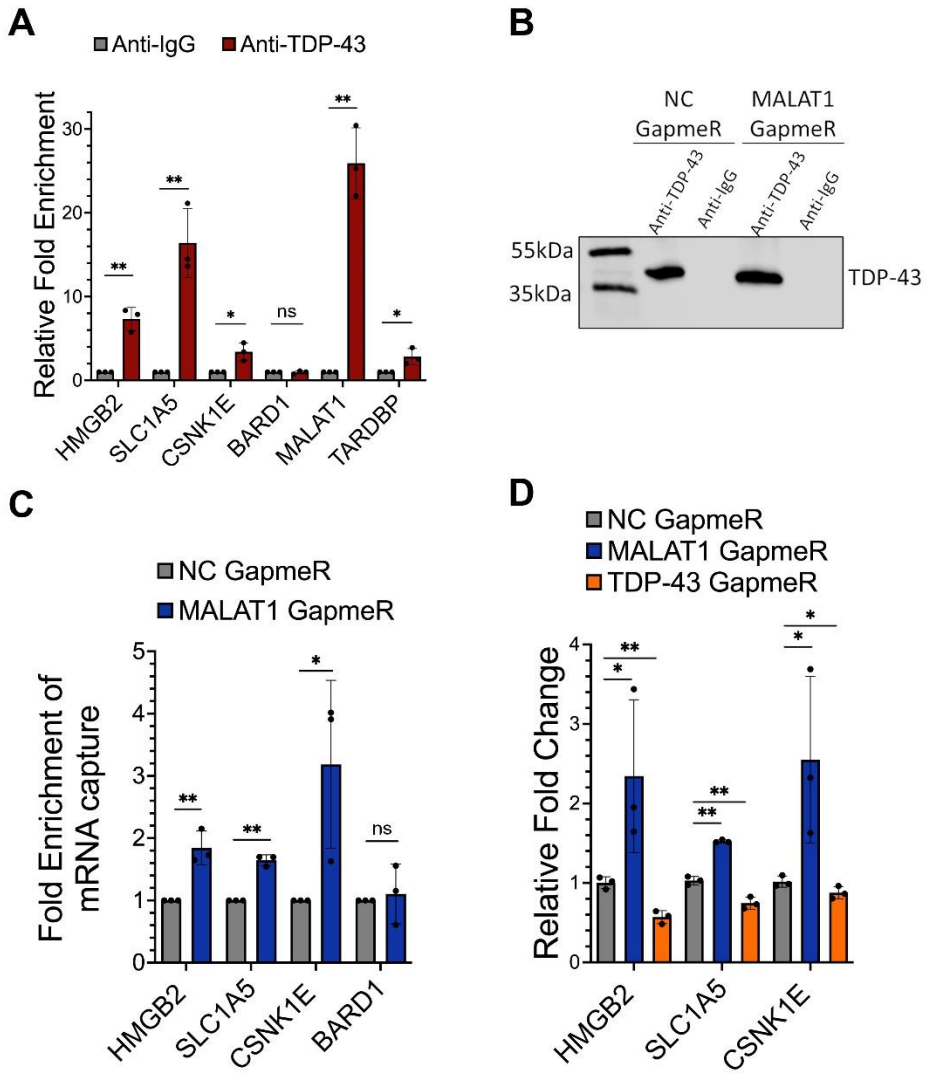


Figure 2.5 - Reduced level of MALAT1 RNA results in relocalization of TDP-43 binding to messenger RNA transcripts in HEK293 cells.

(A) IP-qPCR analysis of TDP-43 enrichment of MALAT1 RNA, TARDBP control, BARD1 control, and 3' UTR of messenger RNA transcripts compared to IgG enrichment of each transcript. (B) Representative Western blot of TDP-43 protein recovery after immunoprecipitation with TDP-43 or IgG control antibody, from negative control (NC) GapmeR or MALAT1 GapmeR treated cells. (C) IP-qPCR analysis of RNA quantification for TDP-43 enrichment of each RNA transcript after knockdown of MALAT1 RNA, compared to TDP-43 enrichment of each RNA transcript in NC GapmeR treated cells. (D) Gene expression changes in 3' UTR target genes after knockdown of MALAT1 RNA or TDP-43 protein, compared to negative control (NC) GapmeR treatment. N=3 for qPCR trials. All experiments in this figure were performed in HEK293 cells. N=3 for RIP experiments. T tests are conducted with two tailed unpaired equal variance conditions. * = $p<0.05$, ** = $p<0.01$. All data are plotted with standard deviation.

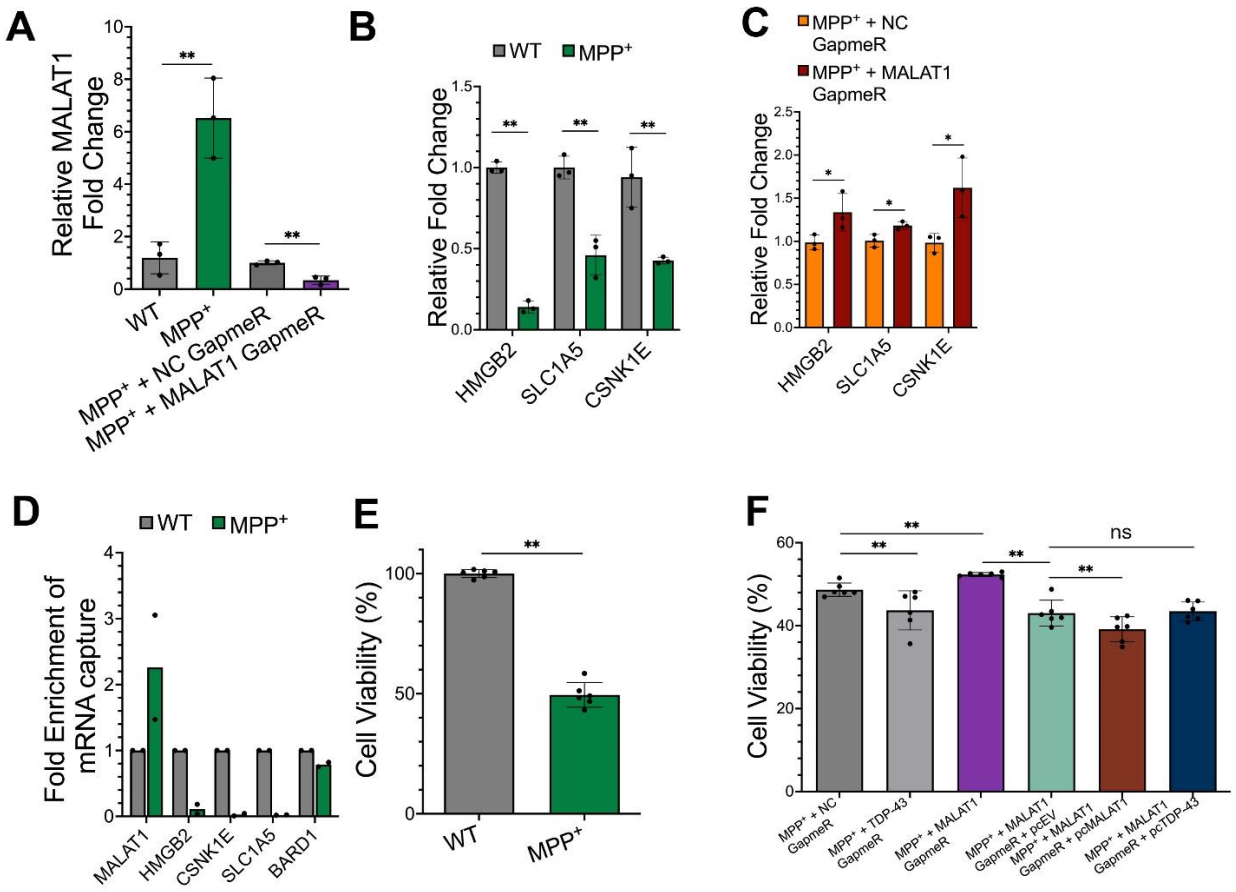


Figure 2.6 - Depletion of MALAT1 RNA protects against MPP+ toxicity in a cell culture model of neurodegeneration.

(A) Quantification of MALAT1 expression level in WT SH-SY5Y, MPP+ treated SH-SY5Y, and MPP+ cells treated with either negative control (NC) GapmeR or MALAT1 GapmeR, using qPCR. (B) Relative expression of mRNA transcripts after MPP+ treatment, compared to expression levels in untreated (WT) SH-SY5Y cells. (C) Relative expression of mRNA transcripts in MPP+ condition after treatment with negative control (NC) GapmeR or MALAT1 GapmeR. (D) IP-qPCR for TDP-43 binding to RNA transcripts in MPP+ treated cells compared to untreated (WT) SH-SY5Y cells. N=2 (E) Viability of SH-SY5Y cells after MPP+ treatment. (F) Viability of SH-SY5Y cells after treatment with the indicated combinations of MPP+, TDP-43 GapmeR, MALAT1 Gapmer, pcEV empty vector control, pcMALAT1 overexpression construct, or pcTDP-43 overexpression construct. All experiments in this figure were performed in SH-SY5Y cells. N=3 for qPCR experiments. N=9 for cell viability experiments. T tests are conducted with two tailed unpaired equal variance conditions. * = $p<0.05$, ** = $p<0.01$ All data are plotted with standard deviation.

Chapter 3 - MALAT1 non-coding RNA regulates alternative splicing of pre-mRNAs through sequence-specific RNA-RNA and RNA-protein interactions

Abstract

Messenger RNAs are subject to multiple layers of gene expression regulation, enabling the creation of a large diversity of RNA transcripts and mature proteins from a smaller number of genes encoded in the human genome. Non-coding RNAs can contribute to alternative splicing (AS) regulation of messenger RNAs, but their mechanisms of action remain unclear in many cases. We find that the broadly expressed MALAT1 non-coding RNA regulates AS of the spermine/spermidine acetyltransferase 1 (SAT1) pre-mRNA through direct, sequence-specific RNA-RNA and RNA-protein binding interactions. Regulation of SAT1 exon X inclusion is required to maintain appropriate levels of polyamine production in human cells, and disrupted SAT1 expression has been associated with negative health outcomes including neurodegeneration and aging. Direct binding of MALAT1 RNA to TDP-43 and SAT1 pre-mRNA enhances AS regulation of SAT1 and promotes cell survival. This novel regulatory mechanism of RNA-RNA and RNA-protein splicing regulation is not limited to the interaction between MALAT1, TDP-43, and SAT1. AS of the liprin-α3 (PPFIA3) exon 16 by cleavage stimulation factor subunit 2 (CSTF2) is similarly enhanced by direct and sequence-specific MALAT1 RNA-RNA and RNA-protein binding interactions with both liprin-α3 pre-mRNA and CSTF2 protein. We propose that the abundant MALAT1 non-coding RNA affects human gene expression through modular RNA-RNA and RNA-protein interaction regions, thus regulating the levels of alternatively spliced mRNA transcripts relevant for human aging and neurodegenerative processes.

Introduction

Human messenger RNA (mRNA) transcripts are modified by multiple post-transcriptional processing steps to ensure appropriate gene expression in specific cell types. Disruption of post-transcriptional processing of mRNAs occurs frequently in human diseases such as cancer and neurodegeneration [1,2]. A key post-transcriptional regulation mechanism for mRNAs in human cells is alternative splicing (AS), which increases genetic complexity and flexibility by allowing multiple transcript isoforms to be encoded at the same gene locus [3]. AS involves the selection of specific exons for inclusion or exclusion into a mature spliced transcript, with unselected introns and exons being removed and degraded. The spliceosome, a large macromolecular structure comprised of splicing proteins and small nuclear RNAs, catalyzes the splicing of unprocessed pre-mRNA transcripts [4,5]. Cell-type specific expression of mRNA isoforms and AS activity of the spliceosome can be altered through interactions with splicing repressor or activator proteins [6]. Additionally, a role for non-coding RNA in splice site selection has been proposed [7,8]. To date, the complex mechanisms of AS regulation of mRNA isoforms in human cells have not been fully elucidated.

AS contributes to the maintenance of transcript regulation at homeostatic levels in healthy and actively growing cells. Production of alternative mRNA isoforms can lead to the formation of truncated protein products with cellular toxicity or aberrant function [9]. Alternative splicing coupled with nonsense-mediated decay (AS-NMD) can be used as a post-transcriptional regulatory mechanism to reduce the levels of enzymes or other functional proteins in response to environmental or metabolic signals [10,11]. This mechanism contributes to the maintenance of steady-state levels of cell-viability regulating genes [12]. The RNA transcript encoding Spermine/Spermidine Acetyltransferase 1 (SAT1, or SSAT RNA) produced from the *sat1* gene is regulated by at least two mechanisms. First, an increase in polyamine concentrations results in

polyamine binding to upstream polyamine responsive element regions to promote overall transcriptional processivity [13, 14]. Second, SAT1 can undergo AS to produce a nonfunctional transcript through inclusion of a poison exon, exon X, which incorporates a premature stop codon in the SAT1-X transcript [15-18]. The SAT1-X transcript is subject to AS-NMD, reducing overall SAT1 protein expression levels to maintain polyamine homeostasis.

SAT1 catalyzes the addition of an acetyl group to polyamines, including spermine and spermidine, to enable their interconversion to other polyamines or their subsequent degradation [19]. Polyamines are abundant organic polycations required for human cell growth [20, 21]. The steady-state level of polyamines in each cell is maintained through the action of several proteins, with SAT1 acting as a key regulatory enzyme. Polyamines promote longevity and neuronal function, with increased dietary polyamines conferring protection against memory loss [22, 23]. Polyamine levels are altered in brain tissue from patients with Parkinson's Disease, and changes in N^1-acetylspermidine, N^8-acetylspermidine, and N^1,N^8-diacetylspermidine levels are associated with cognitive decline [24, 25]. Acetylated versions of spermine and spermidine are directly correlated with SAT1 levels, and SAT1 homeostasis is required for cell survival. The detailed mechanisms of AS choice during regulation of SAT1 and SAT1-X mRNA transcript production are still unknown.

We identified new interactions between the non-coding RNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) and TAR DNA binding protein 43 (TDP-43) that affect AS of SAT1 mRNA. Previous work has implicated both MALAT1 and TDP-43 as splicing regulators of SAT1 [26, 27]. MALAT1 is an 8.5 kb non-coding RNA that is predominately retained in the nucleus and may affect gene expression, alternative splicing, and polyadenylation of mRNAs [28, 29]. MALAT1 dysfunction has been linked to the progression of

neurodegeneration and is proposed to contribute to PD pathogenesis [30, 31]. Aberrant MALAT1 expression is observed in PD-like mouse and human cellular models [32-34]. TDP-43 is a multifunctional nucleic acid binding protein involved in mRNA expression regulation, AS, and mRNA stabilization [35, 36]. TDP-43 has also been proposed to participate in multiple aspects of complex gene expression regulation processes including RNA binding, DNA binding, transcriptional regulation, sub-cellular organization of phase separation and RNA-protein condensate formation (37-42).

Several lines of evidence suggest a functional interaction between MALAT1 and TDP-43 that may be relevant in human disease. Mass spectrometry experiments have demonstrated that MALAT1 interacts with TDP-43 and other splicing factors (43, 44). TDP-43 dysfunction correlates with progression of neurodegenerative disease in brain tissues of patients with Parkinson's Disease [45-47] and MALAT1 shows one of the highest increases in TDP-43 binding in diseased brain tissue compared to normal patient brain tissue [40]. Here, we identify the molecular determinants of a direct three-way MALAT1, TDP-43, and SAT1 pre-mRNA interaction. We evaluate the contribution of these interactions to SAT1 splicing activity and cell survival in a human cell culture model of neurodegeneration. Finally, our results suggest a novel mechanism for regulation of alternative splicing of pre-mRNAs, mediated by sequence-specific RNA-RNA and RNA-protein interactions with MALAT1 non-coding RNA.

Results

MALAT1 and TDP-43 regulate alternative splicing of SAT1 pre-mRNA

The protein-coding mRNA isoform of SAT1 (NM_002970.4;) differs from the unproductive RNA isoform SAT1-X (NR_027783.3) through the exclusion of exon X, located

between exons 3 and 4. A splicing assay for SAT1 mRNA, using primers annealing to exons 3 and 4, was used to measure the splice inclusion and exclusion products for SAT1 compared to SAT1-X under various conditions. Primers complementary to exons 2 and 3 were designed to measure total SAT1 and SAT1-X mRNA levels, and a relative "percentage of splice inclusion" (PSI) value was calculated for SAT1 exon X under different treatment conditions (Fig. 1A).

Alternative splicing (AS) of SAT1 transcripts was measured after knockdown of specific targets with MALAT1 antisense oligonucleotide (ASO) GapmeR, TDP-43 ASO GapmeR, or a negative control random (NC) GapmeR. The ASO GapmeR knockdown method was chosen over siRNA or genetic mutation for the increased ability of ASO GapmeRs to target nuclear RNA transcripts without affecting enhancer or gene regulatory element function at the DNA loci. Knockdown efficiency of MALAT1 and TDP-43 transcripts using ASO GapmeRs ranged from 80-95% reduction of the target RNA (Supplement Fig. 1A). We found that knockdown of MALAT1 and TDP-43 RNA levels reduced the inclusion of SAT1 exon X and lowered the PSI, thereby decreasing levels of the SAT1-X unproductive transcript (Fig. 1B). Double knockdown of MALAT1 and TDP-43 in HEK293 cells further decreased SAT1 exon X PSI compared to individual knockdowns, indicating that MALAT1 and TDP-43 both contribute to SAT1-X production. Next, we evaluated the total mRNA levels of SAT1 transcripts after MALAT1 or TDP-43 perturbation. Compared to the control GAPDH mRNA, total SAT1 and SAT1-X mRNA expression level increased in both the MALAT1 and TDP-43 knockdown conditions (Fig. 1C). Knockdown of an unrelated TDP-43 binding mRNA, HMGB2, using a specific ASO GapmeR had no effect on SAT1 and SAT1-X mRNA expression, validating the specificity of MALAT1 RNA in regulating splicing of SAT1 (Supplement Fig. 1B). In summary, loss of MALAT1 RNA,

TDP-43 protein, or both, results in a decreased level of the spliced exon X transcript (SAT1-X) and an increased level of SAT1 spliced isoform.

TDP-43, MALAT1, and SAT1 form high molecular weight complexes in cells

Given that SAT1 and MALAT1 can both interact with TDP-43, we next investigated whether native TDP-43, MALAT1 and SAT1 interact and form high molecular weight complexes in living cells. We lysed cells and separated endogenous RNA-protein complexes using ultracentrifugation through a 5 to 50% sucrose gradient and measured the distribution of native TDP-43 and RNA transcripts in varying molecular weight complexes (Fig. 2A). As a positive control, total RNA digestion of cellular lysate with RNase resulted in a re-distribution of TDP-43 to lower density fractions of the gradient, representing a loss of higher molecular weight (MW) TDP-43 and RNA complexes (Fig. 2B). These losses were primarily observed in fractions 12 and beyond, which correspond to a sucrose density of 24.8% - 50%. This is the density range in which polysomal fractions can be found, and the loss of TDP-43 complexes in these regions may be due to the inability of ribosome complex formation due to 18S and 28S RNA degradation through RNase [48]. As a negative control, TDP-43 distribution was measured in wild-type and NC GapmeR treated cells. No significant difference in the TDP-43 distribution was observed in the control sample compared to the wild-type sample in any part of the gradient, indicating that treatment with a non-specific GapmeR does not generally affect TDP-43 interactions with other RNA or proteins in the cell (Fig 2C). Next, we compared TDP-43 distribution in control NC GapmeR and MALAT1 GapmeR knockdown conditions in the top 15 fractions where the largest quantities of TDP-43 were observed. We found that MALAT1 knockdown caused a significant re-localization of TDP-43 from higher MW fractions to lower MW fractions. (Fig. 2D). The observed shift occurred mainly between fractions 6 to 12. A sucrose gradient comparison was

also performed between MALAT1 GapmeR conditions and total RNA digestion. There was a clear loss in high MW fraction species in the total RNA digestion fractions (fractions 10-25), which indicates that TDP-43 interacts with other RNAs apart from MALAT1, a result which is consistent with previous reports that TDP-43 binds to thousands of transcripts (Supplement Fig. 2).

Next, we analyzed the fractions collected from the NC GapmeR conditions for RNA distribution through qPCR to determine whether all three components, MALAT1, SAT1, and TDP-43 were present in the same fractions from the gradient. We were able to identify a co-localization of MALAT1 and SAT1 in the sucrose gradient, and show that the two RNAs are localized primarily in fractions 6 to 12, which is where we observe a loss in TDP-43 protein in the MALAT1 knockdown condition (Fig. 2E). We then proceeded to determine if the fractions that exhibited loss in TDP-43 also showed loss in SAT1 mRNA. We compared SAT1 mRNA distribution in the sucrose gradient between NC and MALAT1 GapmeR conditions. We see a marked re-localization in SAT1 mRNA to lower MW fractions in MALAT1 GapmeR conditions, indicating that, similar to TDP-43 distribution, SAT1 mRNA distribution is also MALAT1 dependent (Fig. 2F). A loss in SAT1 mRNA localization occurs in fractions 10-12, which supports the hypothesis that a three-way interaction between SAT1, TDP-43, and MALAT1 is mediated by the presence of MALAT1. The negative control BARD1 mRNA distribution was evaluated to determine whether these effects were specific to SAT1. Upon MALAT1 knockdown, there is no clear re-localization of BARD1 mRNA in the gradient, which we believe is due to its inability to directly bind to MALAT1 or TDP-43 (Fig. 2G).

TDP-43 binds directly to SAT1 and MALAT1 RNA transcripts

We next investigated the binding regions for interaction between TDP-43 and the RNA transcripts. Analysis of TDP-43 cross-linked immunoprecipitation with sequencing (CLIP-seq) data identified direct binding sites on SAT1 and MALAT1 RNA transcripts [49]. These binding sites contain repeat UG-RNA sequences, which are the known consensus TDP-43 binding motif. [40-42, 51] A peak indicating TDP-43 binding was identified on the SAT1 pre-mRNA in the region of intron 4, downstream of the alternatively spliced exon X (Fig. 3A). The predominant TDP-43 CLIP-seq peak on MALAT1 was identified as the 300-nucleotide region from nucleotides 6638-6938 ($MALAT1_{[6638-6938]}$) (Fig. 3B) by comparing peaks across multiple datasets from different cell lines (Supplement Fig. 3A). BARD1 mRNA was used as a negative control because this transcript is expressed but does not have peaks for TDP-43 binding in any of the CLIP-seq datasets. MADD was used as a control for TDP-43 binding because it has been identified in previous studies as a TDP-43 bound gene at intronic regions [27] (Fig. 3C).

TDP-43/SAT1 and TDP-43/MALAT1 binding interactions were confirmed using RNA immunoprecipitation (RNA-IP) with a TDP-43 specific antibody. MALAT1 and SAT1 RNA transcripts were enriched significantly in TDP-43 captures over IgG antibody control captures (Fig. 3D). TDP-43 binding and enrichment in the RNA-IP experiment was validated using the previously characterized autoregulatory site in the 3' UTR of the TARDBP/TDP-43 mRNA [50] and MADD mRNA as positive controls, and BARD1 mRNA as negative control. While similar quantities of TDP-43 protein were enriched in the RNA-IP from MALAT1 and control GapmeR knockdown cells, significant loss in TDP-43 binding to SAT1 pre-mRNA was observed in the MALAT1 knockdown conditions (Fig. 3E). To validate the specificity of this result for SAT1, we analyzed elution levels of another TDP-43 binding pre-mRNA MADD, and observed no significant difference in enrichment in TDP-43 RNA-IP from MALAT1 knockdown cells.

MADD alternative splicing has also previously been shown to be regulated by TDP-43 [27] with the TDP-43 binding site on MADD located between exon 30 and 31, as detailed in Figure 3C. We also evaluated changes in relative PSI of MADD mRNA and found no significant difference in MADD PSI in MALAT1 knockdown conditions, but observed significant increase in MADD PSI in TDP-43 knockdown conditions, as observed in the literature (Supplement Fig. 3B). These results indicate that MADD1 is only a TDP-43 splicing target and not affected by MALAT1. These results suggest that both MALAT1 and TDP-43 are required to bind to SAT1 pre-mRNA and form a three-way interaction.

To eliminate the possibility that loss of TDP-43 and SAT1 mRNA binding was due to a change in cellular localization of SAT1 pre-mRNA after MALAT1 knockdown, we performed fractionation experiments to examine the localization of SAT1. The U1 snRNA and GAPDH mRNA were used as subcellular localization controls to validate the purification of nuclear and cytoplasmic fractions, respectively. No significant difference in SAT1 localization between nucleus and cytoplasm was observed after MALAT1 knockdown (Supplement Fig. 3C).

MALAT1 non-coding RNA has specific binding sites for TDP-43 protein and SAT1 mRNA

To further investigate the molecular mechanism underlying the MALAT1, TDP-43, and SAT1 interaction, we identified regions of MALAT1 RNA that bind to TDP-43 and SAT1. As expected, TDP-43 CLIP-seq peaks aligned with the location of tandem $(UG)_2$ repeat sequences in the binding region of $MALAT1_{[6638\text{-}6938]}$ (Fig. 4A).

Next, we identified MALAT1/SAT1 sequence-specific RNA-RNA interactions within $MALAT1_{[6638\text{-}6938]}$ that could explain the effects of MALAT1 on SAT1 AS. The Inta-RNA computational program [52] was used to perform a Watson-Crick based binding query across the

MALAT1 and SAT1 RNA sequences, and identified a 40-50 nucleotide region of SAT1 and MALAT1 that exhibited RNA-RNA hybridization (Fig. 4B). We then analyzed the MALAT1 and SAT1 RNA sequences to determine RNA nucleotide secondary structure folding energies through the RNAup computational program [53, 54] (Supplement Fig. 4A). No consistent energy profile and/or Watson Crick based RNA-RNA hybridization were observed in a similar analysis of MALAT1 and the negative control mRNA BARD1 (Supplement Fig. 4B).

We quantified the *in vitro* binding affinity of TDP-43 tandem RRMs to RNA targets identified by CLIP-seq using a multiplexed electrophoretic mobility shift assay (mEMSA) [55]. RNA targets for TDP-43 RRM binding were the $MALAT1_{[6638\text{-}6938]}$ sequence, a randomly shuffled version of this sequence termed $MALAT1_{[6638\text{-}shuffled]}$, or a randomly generated 300 nucleotide sequence containing equal ratios of each nucleotide termed $Random_{[300nt]}$ (Fig. 4C). The sequence identity for each construct is detailed in Table S1. We observed strong TDP-43 binding to the $MALAT1_{[6638\text{-}6938]}$ sequence. The shuffled version of this sequence, $MALAT1_{[6638\text{-}shuffled]}$, had lower binding affinity for TDP-43. The completely randomized sequence showed no binding to TDP-43 in the EMSA assay (Fig. 4D). A detailed list of calculated k_d values are shown in Table 1. This confirms the specificity of TDP-43 binding to the $MALAT_{[6638\text{-}6938]}$ sequence. We also performed Coomassie staining on purified TDP-43 tandem RRM constructs to validate our protein purification, and observed single bands of increasing densitometry with increasing protein concentration (Supplement Fig 4C).

$MALAT1_{[6638\text{-}6938]}$ binding to SAT1 mediates efficient SAT1 alternative splicing

To validate that TDP-43-MALAT1 binding was required for efficient SAT1 AS, we performed splicing assays using $MALAT1_{[6638\text{-}6938]}$ and $MALAT1_{[6638\text{-}shuffled]}$ RNA sequences. Both constructs were individually overexpressed in cells with MALAT1 GapmeR knockdown,

and SAT1 PSI ratios were evaluated. Overexpression of the TDP-43 binding region MALAT1$_{[6638-6938]}$ was able to partially rescue SAT1 PSI after MALAT1 GapmeR knockdown (Fig. 5A). The rescue effect was significantly decreased when the shuffled RNA sequence transcript MALAT1$_{[6638-shuffled]}$ was overexpressed, suggesting that the TDP-43-MALAT1$_{[6638-6938]}$ binding sequence identity plays a role in efficient SAT1 AS.

We then deleted the MALAT1-SAT1 hybridizing region in the MALAT1$_{[6638-6938]}$ to create MALAT1$_{[6648-6938-DEL]}$. Compared to the MALAT1$_{[6638-6938]}$ RNA, the MALAT1$_{[6638-6938-DEL]}$ sequence showed decreased efficacy of SAT1 PSI rescue. (Fig. 5B) We ensured that the deletion of the hybridizing sequence on MALAT1$_{[6638-6938]}$ did not affect its ability to bind TDP-43. Multiplex EMSAs were performed with three 100nt fragments – MALAT1$_{[6638-6738]}$, MALAT1$_{[6678-6778]}$, and MALAT1$_{[6763-6863]}$. All fragments showed comparable k_d values and all fragments bound TDP-43 weaker than the whole length MALAT1$_{[6638-6938]}$. A full list of k_d values is provided in Table S2. Given that MALAT1$_{[6638-6738]}$ and MALAT1$_{[6678-6778]}$ both overlap the hybridization region MALAT1$_{[6689-6731]}$ and show comparable k_d to a non-overlapping sequence, the RNA-RNA hybridization sequence is not vital for MALAT1-TDP-43 interaction, and is only involved in MALAT1-SAT1 interaction.

We also observed a successful partial rescue of total SAT1 mRNA expression levels with the addition of the MALAT1$_{[6638-6938]}$ sequence after MALAT1 GapmeR knockdown compared to a empty vector control (Fig. 5C). A partial rescue was also observed with the ectopic overexpression of the MALAT1$_{[6638-6938-DEL]}$ sequence, and no significant rescue was observed upon addition of the MALAT1$_{[6638-6938-shuffled]}$ sequence. The observed changes in SAT1 mRNA expression are consistent with the observed changes in relative PSI.

SAT1 binding to MALAT1 and TDP-43 occurs through sequence-specific interactions

The nucleotide sequence of the SAT1-in4 binding region identified through TDP-43 CLIP-seq peaks contains one consensus $UG_{(4)}$ and two $UG_{(3)}$ repeat sequences, highlighting the preferential TDP-43 binding motif. (Fig. 6A) Additionally, the Inta-RNA hybridization prediction showcases the nucleotide sequence on SAT1-in4 that hybridizes to $MALAT1_{[6638\text{-}6938]}$, as previously shown in Figure 4B. To investigate the importance of the sequence specific interactions between MALAT1 and SAT1, mutant SAT1 transcripts were created and evaluated for their contributions to splicing. Sequences were cloned from HEK293 genomic DNA and cloned into a puc57 vector downstream of a T7 promoter for in-vitro transcription. The base construct [SAT1-in4] is a 437nt RNA in the intron 4 region of SAT1 pre-mRNA that covers the three UG repeat sequences as well as the MALAT1-SAT1 hybridizing sequence at the 3' end. Three additional sequences were generated – the SAT1-in4 sequence with a substitution of the UG repeats to AC [SAT1-in4-ΔUG], the SAT1-in4 sequence with a deletion of the MALAT1 hybridizing region [SAT1-in4-ΔM], and a final SAT1 construct with both UG repeat substitutions and MALAT1-hybridizing sequence deletion [SAT1-in4-ΔUG-ΔM] (Fig. 6B). The full sequences of all constructs have been fully listed in Table S3.

We investigated if these constructs were able to function as we expected. We hypothesized that in-cell, the MALAT1 lncRNA binds to TDP-43 and directs it to SAT1-premRNA, where TDP-43 then interacts with the pre-mRNA to begin the AS process. As such, we designed a competitive EMSA assay to determine if these in-vitro synthesized constructs were able to displace the fluorescently labelled $MALAT1_{[6638\text{-}6938]}$ RNA bound to TDP-43. Increasing molar equivalents of competitor RNA were titrated into the bound TDP-43-$MALAT1_{[6638\text{-}6938]}$ complex and a native gel was used to determine if there had been any displacement of the MALAT1 RNA by the competitor RNA. As expected, the SAT1-in4-ΔUG

and SAT1-in4-ΔUG-ΔM constructs were not able to outcompete the bound MALAT1$_{[6638-6938]}$ due to a lack of UG repeat motifs (Fig. 6C). This re-affirms the statement that UG repeats are needed for TDP-43-RNA binding interactions and is validated by the lack of out competition observed from the BARD1 in-vitro construct as well.

After confirming the importance of UG repeats for RNA-protein interactions, we then decided to investigate if the MALAT1-SAT1 hybridization was necessary for out competition. Competitive EMSA assays were performed for the SAT1-in4 and SAT1-in4-ΔM constructs against the bound TDP-43-MALAT1$_{[6638-6938]}$ complex. Here, we observed that the molar equivalents required for EC_{50} was lower for the SAT1-in4 construct than the SAT1-in4-ΔM (Fig. 6C). We hypothesized that the ability for the two RNA constructs to hybridize allows for the competitor RNA to outcompete at a faster rate. A graphical representation of triplicates is shown for the competitive EMSA, along with visual representations of the gels. EC_{25}, EC_{50} and EC_{75} values are shown in Table 2.

MALAT1 enhances binding of TDP-43 to SAT1 pre-mRNA

We wanted to verify that this difference in competition between SAT1-in4 and SAT1-in4-ΔM was solely due to the RNA-RNA interactions, and not due to any changes in RNA-protein binding affinity between TDP-43 and mutant SAT1 constructs. A quantitative multiplexed EMSA with the SAT1-in4, SAT1-in4-ΔM, and MALAT1$_{[6638-6938]}$ constructs showed no significant difference in binding affinity ($k_{d, EMSA}$) of TDP-43 for the various SAT1 constructs (Fig. 7A) The calculated k_d values and standard deviation are shown in Table 3. We then calculated a binding affinity for the SAT1-in4 and SAT1-in4-ΔM constructs when acting as a competitor ($k_{d, comp}$) using the Lin and Riggs equation [56, 57] and observed a significant increase in binding affinity for the SAT1-in4 construct (Fig. 7B) compared to the SAT1-in4-ΔM

construct. This led us to believe that the difference in out-competition activity was due to a novel RNA-RNA interaction that exists between MALAT1 and SAT1 RNA, which affects binding activity in competitive conditions (due to higher RNA concentrations relative to protein concentrations) and does not affect binding activity in multiplexed EMSA conditions (due to trace concentrations of RNA and excess protein concentrations).

We then validated this proposed RNA-RNA hybridization through a RNA-EMSA gel. All synthesized RNA fragments were run at 10X concentrations on a native gel with 1X concentration of fluorescently labelled $MALAT1_{[6638\text{-}6938]}$. RNA-RNA hybridization was observed as a higher MW band in lanes that contain a SAT1 RNA construct with the MALAT1 hybridization sequence. No RNA-RNA hybridization was observed in lanes with sequence deletion, or negative control BARD1 (Supplement Fig. 5A**)**. We also identified strong energy hybridization between murine Malat1 and SAT1, suggesting conservation of the functional MALAT1-SAT1 interaction in mouse and human (Supplement Fig. 5B). The requirement for UG repeats on the SAT1 pre-mRNA sequence was assessed by comparing the genomic DNA sequence of mouse SAT1 and human SAT1. Two of the three UG repeats are conserved between mouse and human sequences (Supplement Fig. 5C). Evolutionary conservation of UG repeats on SAT1 suggests and important role for TDP-43 binding in SAT1 regulation.

Next, we performed *in vitro* binding experiments with TDP-43-$MALAT1_{[6638\text{-}6938]}$ and the calculated EC_{50} concentrations of the SAT1-in4 or SAT1-in4-ΔM constructs. Samples were incubated with 90-mer biotin-labelled ssDNA probes complementary to the full-length MALAT1 sequence and captured on streptavidin beads. The quantities of captured MALAT1 and SAT1 were measured after RNA elution (Fig. 7C). In both conditions, we eluted a similar amount of $MALAT1_{[6638\text{-}6938]}$, confirming the validity of the experiment. However, there was a significant

loss of SAT1-in4-ΔM elution compared to SAT1-in4 elution (Fig. 7D). Using the MALAT1 hybridization probes ensures that all elutions are RNA species that were directly interacting with the MALAT1 RNA and not mediated by the TDP-43 protein**.** Based on these results, $MALAT1_{[6638\text{-}6938]}$ RNA mediates the interaction between TDP-43 and SAT1. As a negative control, MALAT1 RNA and the control BARD1 RNA were used for a similar capture experiment. There was no elution of BARD1 RNA, indicating that RNA-RNA hybridization is needed for RNA capture and enrichment (Supp. Fig. 5D).

MALAT1 and TDP-43 knockdown affect survival of SH-SY5Y cells

Next, we evaluated whether MALAT1 and TDP-43 mediated AS contributes to cell survival in a disease-relevant cell line. SH-SY5Y cells have previously been used as a model line for Parkinsons Disease research. We validated the GapmeR efficiency for MALAT1 and TDP-43 in SH-SY5Y cells, and used 48 hour transfection time period for optimal RNA knockdown (Supp. 6A, B). We observed a similar change in SAT1 PSI in MALAT1 and TDP-43 GapmeR knockdown conditions in SH-SY5Y cells, indicating that the SAT1 AS mechanism occurs in multiple human cell lines (Fig. 8A). We then performed ectopic overexpression of MALAT1 Ectopic overexpression of MALAT1 or TDP-43 both independently increase SAT1 PSI, suggesting that a similar mechanism of action is used in both SH-SY5Y and HEK293 cells (Fig. 8B).

SAT1 expression levels can vary rapidly in response to metabolic changes in cells. Polyamines such as spermidine can increase SAT1 expression through transcriptional regulation by binding to upstream promoter regions [13, 14, 16]. Spermidine undergoes acetylation to its N-acetylated form through interaction with SAT1 enzyme. We observed that exogenous addition of spermidine increased cell viability and SAT1 gene expression in SH-SY5Y cells (Fig. 8C). To

determine whether this had phenotypic effects in the cells, we added exogenous spermidine to the MALAT1 and TDP-43 GapmeR treated cells and measured the cell viability. The addition of spermidine was able to partially rescue viability in both MALAT1 and TDP-43 GapmeR conditions, indicating that this pathway is important for cell survival (Fig. 8D).

We next investigated how SAT1 is regulated in neurological disease conditions by using MPP^+ (1-methyl-4-phenylpyridinium) treatment of SH-SY5Y cells. In addition to inducing cell death, MPP^+ is known to stimulate an increase in MALAT1 expression in cells [58-60]. Addition of MPP^+ reagent also increased SAT1 PSI, while subsequent addition of MALAT1 GapmeR conversely reduced SAT1 PSI compared to NC GapmeR (Fig. 8E). We validated gene expression and saw the expected increase in MALAT1 RNA levels in MPP^+ conditions, with a decrease upon addition of the MALAT1 GapmeR. TDP-43 mRNA levels did not change significantly in MPP^+ conditions, indicating that SAT1 PSI changes in these conditions are due to changes in MALAT1 expression. However, we observed a unique increase in SAT1 mRNA expression, and subsequent decrease upon MALAT1 GapmeR addition (Fig. 8F). This contrasts with the changes in SAT1 mRNA that we have observed in wild-type conditions upon MALAT1 expression perturbation. We also observed increased cell viability in MPP^+ conditions upon exogenous spermidine addition, when MALAT1 or TDP-43 levels are perturbed (Fig. 8G). This further supports the theory that SAT1 regulation plays a role in neuroblastoma cell viability.

MALAT1 similarly enhances alternative splicing of liprin-α3 by CSTF2 through sequence-specific interactions

Finally, we investigated whether MALAT1 can regulate the splicing activity of other alternative splicing events. Based on mass spectrometry analyses, MALAT1 interacts with numerous splicing and pre-mRNA processing factors, including the cleavage stimulation factor

subunit 2 (CSTF2). CSTF2 binds to mRNA 3'UTRs to promote 3' end cleavage and polyadenylation and has been identified as an alternative splicing regulator in multiple cell lines [61-63]. A comparison of AS genes identified from RNA-seq in CSTF2-knockdown or MALAT1-knockdown cells revealed liprin-α3, also annotated as PTPRF Interacting protein alpha 3 (PPFIA3) mRNA, as a common AS target [26, 61]. Liprins are a family of synaptic scaffolding protein with functions in worm, fly, mouse, and human neurodevelopment [64-66]. Alternative splicing of the liprin-α3 transcript results in two distinct RNA isoforms [64, 67, 68]. Inclusion of liprin-α3 exon 16 leads to productive mRNA and protein production, while the exclusion of exon 16 results in a frameshift induced NMD of the resulting liprin-α3 RNA transcript (Fig. 9A). Due to its similarity in regulation to SAT1 and involvement in neuronal function, we tested the contributions of MALAT1 RNA to CSTF2-mediated alternative splicing of the liprin-α3 mRNA.

To first validate the direct interaction of CSTF2 with MALAT1 and liprin-α3, we re-analyzed existing CSTF2 CLIP-seq data to identify binding regions on the target RNA [61]. The CSTF2 binding region was identified between nucleotides 5785-6126 on the MALAT1 RNA transcript. We also identified a binding peak for CSTF2 on the liprin-α3 (PPFIA3) transcript within intron 16, about 170 nucleotides downstream of the alternatively spliced exon 16. (Fig. 9B). Next, we determined if the $MALAT1_{[5785-6126]}$ region can undergo RNA-RNA hybridization with the liprin-α3 intron 16 region, much like the interaction between $MALAT1_{[6638-6938]}$ and SAT1 intron 4 region. The IntaRNA program was able to validate a strong RNA-RNA hybridization interaction between these two RNA sequences (Fig. 9C). The existence of direct protein binding along with potential RNA-RNA hybridization indicates that a similar mechanism may exist for PPFIA3 AS regulation by CSTF2 and MALAT1.

We next performed splicing assays to measure liprin-α3 exon 16 PSI in the presence or absence of MALAT1 RNA. MALAT1 knockdown resulted in a significant increase in liprin-α3 PSI (Fig. 6D). To test whether this splicing regulation was dependent upon MALAT1 interaction with CSTF2, we cloned and overexpressed the CSTF2 and liprin-α3 binding region, $MALAT1_{[5785-6126]}$, and observed a partial but statistically significant rescue of liprin-α3 exon 16 PSI. In contrast, overexpression of the TDP-43 and SAT1 binding region $MALAT1_{[6638-6938]}$ was unable to rescue the altered PSI of liprin-α3 in MALAT1 knockdown cells. The cloned CSTF2/liprin-α3 and TDP-43/SAT1 binding regions on MALAT1 showed no overlap in sequence or function. These results indicate that the partial rescue of liprin-α3 exon 16 PSI occurred due to MALAT1-mediated CSTF2 localization to liprin-α3 pre-mRNA, similar to the proposed mechanism of regulation of MALAT1-mediated TDP-43 localization to SAT1 pre-mRNA.

Discussion

Alternative splicing regulation of mammalian pre-mRNA transcripts is a complex process involving multiple factors. Accurate exon or intron inclusion choices are required for cell survival, and dysregulation of gene expression can contribute to disease pathology including in cancer and neurodegeneration. We report a novel mechanism by which MALAT1 non-coding RNA can enhance alternative splicing of pre-mRNA transcripts in human cells. Three-way interactions between MALAT1, target pre-mRNAs, and AS regulatory proteins based on seqeuence-specific interactions can facilitate splicing choice in healthy cells. Disruption of these processes by knockdown of MALAT1 transcript leads to alterations in splicing and cell death in a culture model of human neurodegenerative disease.

Our data show that MALAT1/SAT1, MALAT1/TDP-43, and SAT1/TDP-43 interactions occur to create a functional three-way binding and regulate splicing of SAT1 pre-mRNA. Loss of MALAT1 or TDP-43 results in changes in SAT1 splicing choice. One possible mechanism through which the regulation of target species by MALAT1 could occur is through the formation of an RNA-RNA-protein complex which shifts the balance of TDP-43 binding to SAT1 transcripts in the nucleus and away from other pre-mRNA transcripts (Fig. 10A). Directed transfer of nucleic acid binding has been proposed as a means of regulating protein interactions with multiple RNAs [69]. Another possibility is that MALAT1 binding to SAT1 and TDP-43 affects the local sub-nuclear concentration of protein and mRNA partners. Stable higher-order multimeric complexes of MALAT1, TDP-43, and SAT1 could facilitate prolonged interactions in splicing complexes and increase the time that these transcripts spend in proximity to splicing factors. We found that stable high-molecular weight complexes of TDP-43 and SAT1 in cells were dependent upon the expression of MALAT1 in the nucleus. Future experiments will be needed to characterize the components of these high-molecular weight complexes and the dynamics of the three-way MALAT1/TDP-43/SAT1 interaction.

Notably, we found that the MALAT1/TDP-43/SAT1 interaction is not a unique case, and a similar splicing enhancement was observed for MALAT1/CSTF2/liprin- α3. Three-way binding of MALAT1 to mRNA and protein partners may be used as a general mechanism to rapidly adapt and maintain appropriate splicing balance by coordinating RBP and specific pre-mRNA transcript interactions in the nucleus. How RNA-binding alternative splicing regulator proteins find their targets in the crowded nuclear environment is not well understood, especially because known AS regulators like TDP-43 have thousands of pre-mRNA binding partners. Participation of MALAT1 as an additional RNA component of these complexes could enable

added selectivity in binding to restrict or enhance the frequency of TDP-43 and SAT1 interactions. For example, increased or decreased expression of MALAT1 could shift the splicing balance to the SAT1 mRNA product that is most needed at a given time in the cell, based on metabolic response or other signals that regulate MALAT1 RNA levels. TDP-43 can form phase separated domains in the nucleus, which could also be affected by MALAT1 binding or coordination of pre-mRNA and TDP-43 three-way interactions.

We noted that ectopic overexpression of $MALAT1_{[6638\text{-}6938]}$ in MALAT1 knockdown conditions did not fully rescue SAT1 PSI, potentially because of the contribution of other interactions that are disrupted after MALAT1 depletion and cannot be bound by the TDP-43 interaction region of $MALAT1_{[6638\text{-}6938]}$. Other SF and epigenetic regulators may also be required for SAT1 AS, such as the H3K9 demethylase KDM3A, or Serine-Arginine splicing factor protein SRSF1 [18, 26]. The complexity of AS regulation events may also explain why MALAT1 and TDP-43 knockdowns do not abrogate SAT1 AS completely.

MALAT1 may use distinct binding regions to bring mRNA and protein partners together in the nucleus. Although the human MALAT1 transcript is more than 7500 nucleotides in length, the binding sites for TDP-43 and SAT1 on MALAT1 occur within a single 300-nucleotide region. Similarly, the binding sites for CSTF2 and liprin-α3 occur within a distinct 300-nucleotide region of the MALAT1 RNA transcript (Fig. 10B). A network of RNA-RNA and RNA-protein interactions brought together by MALAT1 binding could enable rapid transcriptional changes through proximity guided splicing.

We also find that perturbation of MALAT1 or TDP-43 affects survival in neuroblastoma cells. Polyamines in the nucleus can counter-balance highly charged RNA species and stabilize RNA structure [70]. Addition of exogenous spermidine promotes cell viability and neuronal

function [71-73]. Treatment of neuroblastoma cells with the small molecule MPP+ induces MALAT1 overexpression, while knockdown of MALAT1 under these conditions is protective. The effects of MALAT1 knockdown in MPP+ treated SH-SY5Y cells were partially complemented by addition of exogenous polyamines, suggesting that changes in SAT1 splicing contribute to the death phenotype observed in these cells**.** Polyamines, MALAT1 RNA, mRNAs, and RNA-binding proteins may participate in an interconnected network of macromolecules contributing to the regulation of cellular homeostasis.

Non-coding RNA coordination of three-way RNA-RNA-protein interactions is a novel mechanism to modify complex RBP-mRNA networks and regulate gene expression at the post-transcriptional level. These findings could lead to new therapeutic approaches to combat neurodegeneration or shift alternative splicing choice for specific pre-mRNAs to affect cell survival.

Tables and Sequences

Table 3.1 – Calculated k_d values of MALAT1$_{[6638\text{-}6938]}$ derived constructs

Construct	**k_d (nM)**
MALAT1$_{[6638\text{-}6938]}$	274.21
MALAT1$_{[6638\text{-}6938\text{-}shuffled]}$	1496.06
Random 300nt	12226.81

Table 3.2 – Calculated EC_{25}, EC_{50} and EC_{75} values for SAT1-in4 derived constructs

Competitor RNA Construct	**EC_{25} (Molar Fold)**	**EC_{50} (Molar Fold)**	**EC_{75} (Molar Fold)**
SAT1-in4	1.39	2.13	3.25
SAT1-in4-Δ**M**	2.13	3.74	6.56
SAT1-in4-ΔUG	39.03	100	256.24
SAT1-in4-ΔUG-Δ**M**	25.76	100	388.18
BARD1	18.07	10	553.54

Table 3.3 – Calculated $K_{d, EMSA}$ and $K_{d, Comp}$ values for SAT1-in4 constructs

	$K_{d, EMSA}$			**Avg $K_{d, EMSA}$**	**$k_{d, Comp}$**			**Avg $K_{d, comp}$**
	1	**2**	**3**		**1**	**2**	**3**	
SAT1-in4	173.2	176.1	197.9	182.4	125.8	109.0	130.0	121.6
SAT1-in4-ΔM	183.2	184.9	192.9	187.0	181.6	157.5	141.5	160.2
$MALAT1_{[6638-6938]}$	153.3	145.9	138.6	145.9				

Table 3.4 - Oligonucleotide primers used for quantitative RT-PCR and relative PSI analysis

Primer Name	**Sequence (5' to 3')**
BARD1 FWD	GCCAAAGCTGTTTGATGGAT
BARD1 REV	CGAACCCTCTCTGGGTGATA
GAPDH FWD	GGGCTCTCCAGAACATCATCC
GAPDH REV	GTCCACCACTGACACGTTGG
HMGB2 FWD	TGCTCTGAACATCGCCCAAA
HMGB2 REV	GCTTCACTTTTGCCCTTGGC
MADD FWD	TCTCAAGCGCCTGGTG
MADD REV	CTCTCGAAGGTCATGCAG
MALAT1 FWD	GAAGGAAGGAGCGCTAACGA
MALAT1 REV	TACCAACCACTCGCTTTCCC
$MALAT1_{[6638\text{-}6938]}$ FWD	ATCCCGCTGCTATTAGAATG
$MALAT1_{[6638\text{-}6938]}$ REV	GCTTGTTTGGAATGTTTCTT
PSI MADD FWD	GACCTGAATTGGGTGGCGAGTTCCCT
PSI MADD REV	CATTGGTGTCTTGTACTTGTGGCTC
PSI PPFIA3 FWD	CAGCTGGAGGCCATCAACAAG
PSI PPFIA3 REV	GGAGCTCCAGCTTCCTCCTTAG
PSI SAT1 FWD	TTGCAGAAGTGCCGAAAGAGC
PSI SAT1 REV	TCCAACCCTCTTCACTGGAC
SAT1 FWD	TGTTTATCCGTCACTCGCCG
SAT1 REV	TTTAGCCAGCTCCTTGATCAGC
TARDBP FWD	TCAGGGCCTTTGCCTTTGTT
TARDBP REV	TGCTTAGGTTCGCATTGGAT
TUBULIN FWD	CCAGACAACTTTGTATTTGGTCAGT
TUBULIN REV	CGTACCACATCCAGGACAGAAT
U1 FWD	TTACCTGGCAGGGGAGATAC
U1 REV	TCCCACATTTGGGGAAATC

Methods

Cell culture and transfection

Cell lines for this study were obtained from and verified by the American Type Culture Collection. HEK293 cells (ATCC CRL-1573) were maintained in Eagle's Minimum Essential Medium (EMEM) media, SH-SY5Y neuroblastoma cells (ATCC CRL-2266) were maintained in a 1:1 mixture of EMEM and Gibco Ham's F12 media, and K562 cells (ATCC CCL-243) were maintained in IMDM media. All media were supplemented with 10% fetal bovine serum (FBS) and 1% L-glutamine. Cells were grown at 37 ˚C with 5% CO_2 and routinely tested for mycoplasma contamination. All transfections for knockdown experiments were performed with locked nucleic acid (LNA) GapmeRs, to achieve efficient nuclear targeting of RNA and to avoid the introduction of genetic mutations that could affect enhancer function at the DNA locus. GapmeRs with antisense sequence specific to MALAT1 (LG0000003-DDA) and a non-targeting control GapmeR (LG00000002-DDA) were purchased from Qiagen. GapmeRs for TDP-43 (LG00792378-DDA) and HMGB2 (LG00802781-DDA) were designed using Qiagen GeneGlobe. RNA knockdowns were performed by transient transfection as per manufacturer protocols, using OptiMeM 1 Reduced Serum Medium. HEK293 and K562 cells were transfected with LNA GapmeRs using Lipofectamine 2000 for 24 hours, while SH-SY5Y cells were transfected using Lipofectamine 3000 for 48 hours. Cells were collected and assessed for cell viability after the indicated time period of treatment. Knockdown efficiency for each transcript was verified through RT-qPCR. For complementation experiments, overexpression plasmids were co-transfected with LNA GapmeRs in HEK293 cells with Lipofectamine 2000. Overexpression plasmids were transfected in SH-SY5Y cells with Lipofectamine 3000 and P3000 reagent for 48 hours.

CLIP-seq analysis

Three CLIP datasets were used to observe TDP-43 binding – SRR4044755, ERR9192743, and ERR039855. In addition, the SRR488740 CLIP datatset was used for CSTF2 binding.
The datasets with UMIs were processed with UMI Tools [74]. All datasets were trimmed with Cutadapt [75] to remove adapters, low quality bases, and short sequences (<20 bp). Datasets with barcodes were filtered only for reads with the correct barcode. Reads were aligned to the human genome (hg38) with STAR [76], and the alignment files were indexed with samtools [77]. Graphic representations of CLIP peaks were created using the Integrative Genomics Viewer [78], and figures with CLIP binding regions were created using Biorender.

Multiplexed fluorescent EMSA Assay

mEMSA experiments were performed as previously described [55]. TDP-43 protein RRM concentration was calculated and diluted to a range of concentrations from 1nM to 10μM. RNA transcripts were generated through in vitro transcription using NEB HiScribe T7 High Yield RNA synthesis kit (E2040S). Each RNA transcript was transcribed with the addition of a fluorophore labelled UTP for fluorescence visualization. Fluorescein-12-uridine-5'-triphosphate from Enzo Life Sciences, Inc. (catalog #ENZ-42834), Cyanine 3-uridine-5'-triphosphate (enhanced) from Enzo Life Sciences, Inc. (catalog #ENZ-42505), and cyanine 5-UTP from Apexbio Technology LLC (purchased through Fisher Scientific catalog # 50-199-8343) were chosen to target three unique wavelengths. All RNA fragments were pooled together at a equimolar concentration of 2nM and heat refolded. The RNA fragments were then left to incubate with the different TDP-43 concentrations for 30 minutes at room temperature, in the presence of EMSA binding buffer diluted to 1X concentration. The incubated samples were then run on a 4% TBE polyacrylamide gel for 1 hour, and then imaged on a Typhoon FLA 9500 with

pixel size of 50um for each fluorophore. The lasers and filter combinations were 473nm/BPB1, 532nm/BPG1, and 635 nm/LPR for fluorescein Cy3, and Cy5, respectively. Gel images were quantified with ImageStudioLite and colored using ImageJ for visual representation. The binding affinity, $k_{d,\,EMSA}$, was calculated by fitting to Equation (1) using the nonlinear least squares method in R:

$$(1)\ Fraction\ bound = \frac{[P]^n}{[P]^n + K_d}$$

Competitive EMSA Assay

Competitive EMSA Assays were performed using the protocol devised by Ryder, Recht and Williamson [57] . Fluorescent-labelled $MALAT1_{[6638\text{-}6938]}$ was synthesized along with non-fluorescent competitive RNA fragments. All RNA fragments were diluted to 50nM and heat refolded. Increasing concentrations of competitor RNA were incubated with labelled MALAT1 and TDP-43 RRMs, in the presence of 1X EMSA Binding buffer, for 30 minutes at room temperature. Samples were run on a 4% TBE polyacrylamide gel for 45 minutes, and imaged on the Typhoon FLA 9500. Gel images were quantified with ImageStudioLite and colored using ImageJ for visual representation. The fraction of MALAT1 bound to TDP-43 was calculated through Equation (2):

$$(2)\ Fraction\ bound = \frac{band\ intensity\ of\ bound\ MALAT1}{band\ intensity\ of\ bound\ MALAT1 + unbound\ MALAT1}.$$

Molar equivalents of competitor RNA were calculated for 25%, 50% and 75% bound MALAT1 states. The fraction bound values were plotted alongside the molar excess of competitive RNA using R Studio, for all competitive RNA fragments. The binding affinity, $k_{d,\,comp}$ was calculated by fitting to equation (3) as provided by Lin and Riggs [56, 57].

$$(3)\ K_{d,comp} = \frac{2K_d IC_{50}}{2P - R - 2K_{d,EMSA}}$$

Splicing Assay for percent splice inclusion (PSI)

Splicing assays were performed using PCR primers overlapping the alternatively spliced exon. Samples were processed in biological triplicates and were amplified using the appropriate PCR primers for 35 cycles. The PCR products were visualized on a 8% TBE polyacrylamide gel. All PCR runs included an additional control run with GAPDH primers to ensure amplification and starting sample concentrations were within range of each other. The gels were incubated with SYBR safe DNA stain before visualization, and the band density of inclusion and exclusion PCR products were calculated using ImageStudioLite. PSI was calculated as the density of the inclusion product divided by the total density of the inclusion and exclusion product, and are plotted as relative fold changes. All primers used for PSI calculations are detailed in Table 4.

RNA Immunoprecipitation

RNA immunoprecipitation was performed as previously detailed [82] with a few modifications. HEK293 cells were transfected with either NC or MALAT1 GapmeR for 24 hours. The nuclear extract of HEK-293 cells were then isolated through Gagnon Fractionation [79] with a Hypotonic Lysis Buffer for removing cytoplasmic components (10mM Tris (pH 7.5), 10mM NaCl, 3mM $MgCl_2$, 0.3% (v/v) NP-40 and 10% (v/v) glycerol with 1× EDTA- free Protease Inhibitor Cocktail (PIC)) and a Nuclear Lysis Buffer for isolation of the nuclei components (20 mM Tris pH 7.5, 150 mM KCl, 3 mM $MgCl_2$, 0.3% (v/v) NP-40 and 10% (v/v) glycerol, with 1× PIC). A sample of 20 µL of lysate was saved for RIP input analysis. RNA immunoprecipitation was performed as described in Gagliardi and Matarazzo 2016. The following buffers were prepared for RNA-IP:

NT-2 buffer (5×): 250 mM Tris-HCl pH 7.4, 750 mM NaCl, 5 mM $MgCl_2$, 0.25 % NP-40. NET-2 buffer (1×): 1× NT-2 buffer supplemented with 20 mM EDTA pH 8.0, 1 mM DTT, 200 units/ml RNase OUT.

75µL of Dynabeads Protein-G were incubated with 5µg of antibodies corresponding to TDP-43 (10782-2-AP) and IgG (30000-0-AP) for 1.5 hours at RT. Beads were washed with 1×-NT-2 buffer and the pooled and lysed supernatant were added to each sample with NET-2 buffer and rotated overnight at 4°C. Samples were then washed 6 times with 1× NT-2 buffer and split into equal halves for protein and RNA extraction. Proteins were extracted from the beads through elution with 2X SDS buffer and ran on a 12% PAGE gel to assess immunoprecipitation efficiency. RNA was extracted and purified through ethanol precipitation onto silane-treated magnetic beads. Purified RNA was resuspended in 10 µL UP H_2O and was processed for RT-qPCR.

Quantitative real-time PCR (qPCR) Assays

To quantify RNA levels, qPCR was performed on a Thermo QuantStudio 3 using gene-specific primers and a housekeeping gene targets for controls, using SYBR green for detection and ROX for normalization. Total RNA was isolated from 1 x 10^6 cells using the RNeasy Mini Kit and resuspended in 30 µL of Invitrogen UltraPure H_2O. Next, cDNA was synthesized using 1 µg of RNA and the ProtoScript II Reverse Transcriptase, using the protocol provided by the manufacturer. The knockdown efficiency for different treatments was calculated using the ΔCt method and was measured in technical triplicates for each experimental sample. RIP Fold Change was calculated by determining the % of RNA compared to input sample [80]. All primer sequences used for qPCR are listed in Table 4.

Cell Viability Assays

The Cell Counting Kit-8 (CCK8, Abcam) was used to measure cell viabilities. SH-SY5Y cells were seeded onto a 96 well plate and transfected with LNA GapmeRs. In addition, they were either complemented with exogenous spermidine for 4 hours, or the MPP^+ reagent for 48 hours. 6 μL of CCK8 reagent was added to each well and incubated at 37°C for 2 hours. 96 well plates were then shaken for 30 seconds, followed by imaging at 460 nm using a Synergy HTX multimode microplate reader.

Computational analysis of RNA-RNA Hybridization

RNA-RNA hybridization was computed through two methods – the IntaRNA software, and the RNAup program. For determining Watson-Crick based hybridization, the two RNA sequences were entered on the IntaRNA program. 5 interactions per RNA pair were calculated, with a minimum parameter of 5 base-pairs required for interaction analysis. For minimum free energy hybridization values, the two RNA sequences of interest were uploaded to the RNA web server. For the output energy profile, basic options were standardized as recommended by the program.

Fluorescent-modified RNA Antisense Purification

Heat-refolded Cy2-labelled $MALAT1_{[6638\text{-}6938]}$ was incubated with either Cy3-labelled SAT1-in4 or SAT1in4-ΔM and TDP-43 RRMs in individual tubes for 30 minutes at room temperature. For hybridization, 4ug of biotin-labelled 90mer ssDNA probes targeting the full length of the MALAT1 lncRNA were denatured at 85°C for 3 minutes followed by cooling on ice for 1 minute. The probes were then incubated with the complex tubes and incubated at 37°C for 2

hours with mixing at 1100rpm, with a 15 second on-off cycle After the incubation, 10 μL of Invitrogen MyOne C1 streptavidin beads were added to the samples and mixed for another 30 minutes at 37°C. Hybrids were then transferred to a magnetic rack and isolated, followed by 3 washes using 1X NT-2 buffer. RNA samples were then eluted by Proteinase-K digestion for 1 hour at 50°C, followed by silane cleanup to enrich for the MALAT1-bound RNA. The eluted RNA was then subjected to Reverse Transcription, followed by qPCR for enrichment analysis.

Density ultracentrifugation of high molecular weight complexes

To prepare sucrose gradients for ultracentrifugation, 50mL of 5% (w/v) and 50% (w/v) sucrose buffers were prepared in 20mM Tris-HCl, 150mM KCl, and 1mM MgCl buffer. The sucrose buffers were cooled down to 4°C. For the cellular input, 5 x 10^6 K562 cells were collected and lysed for 15 minutes on ice with 1X lysis buffer (20 mM Tris·HCl, pH 7.5, 150 mM KCl, 1 mM MgCl2, 1 mM DTT, 1 mM PMSF (Protease Inhibitor), 0.2% Triton X-100, 20 units/mL DNase I, 200 units/mL RNAse Inhbitor, and 1X Protease Inhibitor Cocktail). The lysed solution was sonicated thrice at 30% amplitude for 10 sec and incubated for another 15 min on ice. Lysate was centrifuged at 13,000 RPM for 15 min at 4°C, and the supernatant was collected and diluted to 1 mL total volume. The 5% to 50% w/v sucrose gradient was prepared using a Biocomp Gradient Master 108. The cell lysate was loaded onto the gradient and complexes were separated by ultracentrifugation in a Beckman SW-41 TI swinging bucket rotor at 35,000 RPM for 18 hours at 4°C. Individual fractions of 500 μL each were collected by pipetting from the top to the bottom of the gradient and then processed for RNA and protein analysis. For RNA analysis, the fractions were subjected to TRIzol extraction according to the manufacturer's protocol. Collected RNA

fractions were converted to cDNA by reverse transcription for qPCR analysis and quantification. For protein analysis, a 20 μL sample of each collected fraction was diluted with 6X loading buffer and separated by SDS-PAGE for Western Blot analysis.

TDP-43 tandem RRM Purification

The RNA recognition motifs (RRMs) (Q101-Q269) of TDP-43 were purified from a construct with an N terminus 6x-His-Gb1 tag and a TEV cleavage site. This construct was transformed into BL21 *E. coli* and expression was induced with 2% (w/v) L-arabinose at OD600 ~0.7 with in LB medium with ampicillin at 16°C overnight. The fusion protein was purified with a HisTrap FF 1 mL column (Cytiva 17-5319-01) with Buffer A (20 mM HEPES, pH 7.5, 1 M NaCl, 10% glycerol (w/v), 30 mM imidazole and 5 mM 2-mercaptoethanol) and Buffer B (same as Buffer A, but containing 300 mM imidazole) on an AKTA pure FPLC system at 4 °C (Cytiva). The eluted protein was dialyzed overnight at 4 °C in Buffer A, then cleaved with TEV protease [81] for 4 hours at 30 °C. The protein was reapplied to the HisTrap FF column and the flow through containing untagged TDP-43 was collected. The purified protein was concentrated with a 10kDa MWCO centrifugal concentrator (Thermo 88527) and buffer exchanged into 50mM HEPES pH 7.5, 300mM NaCl, 10% glycerol, and 5mM beta-mercaptoethanol before being flash frozen in liquid nitrogen and stored at –80 °C. Increasing concentrations of TDP-43 RRMs were run on a 4% SDS-PAGE gel and post-stained with Coomassie reagent to visualize homogenic purity.

References

1. Corbett A. H. (2018). Post-transcriptional regulation of gene expression and human disease. Current opinion in cell biology, 52, 96–104. https://doi.org/10.1016/j.ceb.2018.02.011

2. Carey, K. T., & Wickramasinghe, V. O. (2018). Regulatory potential of the RNA processing machinery: implications for human disease. Trends in Genetics, 34(4), 279-290.

3. Tao, Y., Zhang, Q., Wang, H., Yang, X. and Mu, H., 2024. Alternative splicing and related RNA binding proteins in human health and disease. Signal Transduction and Targeted Therapy, 9(1), p.26.

4. Yan, C., Wan, R., & Shi, Y. (2019). Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. Cold Spring Harbor perspectives in biology, 11(1), a032409.

5. Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. cell, 136(4), 701-718.

6. Ule, J., & Blencowe, B. J. (2019). Alternative splicing regulatory networks: functions, mechanisms, and evolution. Molecular cell, 76(2), 329-345.

7. Pisignano, G., & Ladomery, M. (2021). Epigenetic regulation of alternative splicing: how LncRNAs tailor the message. Non-coding RNA, 7(1), 21.

8. Bardou, F., Ariel, F., Simpson, C. G., Romero-Barrios, N., Laporte, P., Balzergue, S., ... & Crespi, M. (2014). Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. Developmental cell, 30(2), 166-176.

9. Nikom, D., & Zheng, S. (2023). Alternative splicing in neurodegenerative disease and the promise of RNA therapies. Nature Reviews Neuroscience, 24(8), 457-473.

10. Brogna, S., & Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nature structural & molecular biology, 16(2), 107-113.

11. Tan, K., Stupack, D. G., & Wilkinson, M. F. (2022). Nonsense-mediated RNA decay: an emerging modulator of malignancy. Nature Reviews Cancer, 22(8), 437-451.

12. Neff, A. T., Lee, J. Y., Wilusz, J., Tian, B., & Wilusz, C. J. (2012). Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome research, 22(8), 1457-1467.

13. Wang, Y., Xiao, L., Thiagalingam, A., Nelkin, B. D., & Casero, R. A. (1998). The identification of a cis-element and a trans-acting factor involved in the response to polyamines and polyamine analogues in the regulation of the human spermidine/spermine N 1-acetyltransferase gene transcription. Journal of Biological Chemistry, 273(51), 34623-34630.

14. Wang, Y., & Casero Jr, R. A. (2006). Mammalian polyamine catabolism: a therapeutic target, a pathological problem, or both?. Journal of biochemistry, 139(1), 17-25.

15. Pegg, A. E. (2009). Mammalian polyamine metabolism and function. IUBMB life, 61(9), 880-894.

16. Pegg, A. E. (2008). Spermidine/spermine-N 1-acetyltransferase: a key metabolic regulator. American Journal of Physiology-Endocrinology and Metabolism, 294(6), E995-E1010.

17. Pozzi, B., Bragado, L., Mammi, P., Torti, M. F., Gaioli, N., Gebhard, L. G., ... & Srebrow, A. (2020). Dengue virus targets RBM10 deregulating host cell splicing and innate immune response. Nucleic acids research, 48(12), 6824-6838.

18. Baker, M., Petasny, M., Taqatqa, N., Bentata, M., Kay, G., Engal, E., ... & Salton, M. (2021). KDM3A regulates alternative splicing of cell-cycle genes following DNA damage. Rna, 27(11), 1353-1362.

19. Mandal, S., Mandal, A., Johansson, H. E., Orjalo, A. V., & Park, M. H. (2013). Depletion of cellular polyamines, spermidine and spermine, causes a total arrest in translation and growth in mammalian cells. Proceedings of the National Academy of Sciences, 110(6), 2169-2174.

20. Mandal, S., Mandal, A., & Park, M. H. (2015). Depletion of the polyamines spermidine and spermine by overexpression of spermidine/spermine N 1-acetyltransferase 1 (SAT1) leads to mitochondria-mediated apoptosis in mammalian cells. Biochemical Journal, 468(3), 435-447.

21. Saiki, S., Sasazawa, Y., Fujimaki, M., Kamagata, K., Kaga, N., Taka, H., Li, Y., Souma, S., Hatano, T., Imamichi, Y., Furuya, N., Mori, A., Oji, Y., Ueno, S. I., Nojiri, S., Miura, Y., Ueno, T., Funayama, M., Aoki, S., & Hattori, N. (2019). A metabolic profile of polyamines in parkinson disease: A promising biomarker. Annals of neurology, 86(2), 251–263. https://doi.org/10.1002/ana.25516

22. Gupta VK, Scheunemann L, Eisenberg T, Mertel S, Bhukel A, Koemans TS, Kramer JM, Liu KS, Schroeder S, Stunnenberg HG, Sinner F, Magnes C, Pieber TR, Dipt S, Fiala A, Schenck A, Schwaerzel M, Madeo F, Sigrist SJ. Restoring polyamines protects from age-induced memory impairment in an autophagy-dependent manner. Nat Neurosci. 2013 Oct;16(10):1453-60. doi: 10.1038/nn.3512. Epub 2013 Sep 1. PMID: 23995066.

23. Eisenberg T, Abdellatif M, Schroeder S, Primessnig U, Stekovic S, Pendl T, Harger A, Schipke J, Zimmermann A, Schmidt A, Tong M, Ruckenstuhl C, Dammbrueck C, Gross AS, Herbst V, Magnes C, Trausinger G, Narath S, Meinitzer A, Hu Z, Kirsch A, Eller K, Carmona-Gutierrez D, Büttner S, Pietrocola F, Knittelfelder O, Schrepfer E, Rockenfeller P, Simonini C, Rahn A, Horsch M, Moreth K, Beckers J, Fuchs H, Gailus-Durner V, Neff F, Janik D, Rathkolb B, Rozman J, de Angelis MH, Moustafa T, Haemmerle G, Mayr M, Willeit P, von Frieling-Salewsky M, Pieske B, Scorrano L, Pieber T, Pechlaner R, Willeit J, Sigrist SJ, Linke WA, Mühlfeld C, Sadoshima J, Dengjel J, Kiechl S, Kroemer G, Sedej S, Madeo F. Cardioprotection and lifespan extension by the natural polyamine spermidine. Nat Med. 2016 Dec;22(12):1428-1438. doi: 10.1038/nm.4222. Epub 2016 Nov 14. PMID: 27841876; PMCID: PMC5806691.

24. Vrijsen, S., Houdou, M., Cascalho, A., Eggermont, J., & Vangheluwe, P. (2023). Polyamines in Parkinson's disease: balancing between neurotoxicity and neuroprotection. Annual Review of Biochemistry, 92, 435-464.

25. Lewandowski, N. M., Ju, S., Verbitsky, M., Ross, B., Geddie, M. L., Rockenstein, E., ... & Small, S. A. (2010). Polyamine pathway contributes to the pathogenesis of Parkinson disease. Proceedings of the National Academy of Sciences, 107(39), 16970-16975.

26. Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., ... & Prasanth, K. V. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell, 39(6), 925-938.

27. De Conti, L., Akinyi, M. V., Mendoza-Maldonado, R., Romano, M., Baralle, M., & Buratti, E. (2015). TDP-43 affects splicing profiles and isoform production of genes involved in the apoptotic and mitotic cellular pathways. Nucleic acids research, 43(18), 8990-9005.

28. Arun, G., Aggarwal, D., & Spector, D. L. (2020). MALAT1 long non-coding RNA: functional implications. Non-coding RNA, 6(2), 22.

29. Arun, G., & Spector, D. L. (2019). MALAT1 long non-coding RNA and breast cancer. RNA biology, 16(6), 860-863.

30. Li, L., Xu, Y., Zhao, M., & Gao, Z. (2020). Neuro-protective roles of long non-coding RNA MALAT1 in Alzheimer's disease with the involvement of the microRNA-30b/CNR1 network and the following PI3K/AKT activation. Experimental and molecular pathology, 117, 104545.

31. Abrishamdar, M., Jalali, M. S., & Rashno, M. (2022). MALAT1 lncRNA and Parkinson's Disease: The role in the Pathophysiology and Significance for Diagnostic and Therapeutic Approaches. Molecular neurobiology, 59(9), 5253-5262.

32. Chen, Q., Huang, X., & Li, R. (2018). lncRNA MALAT1/miR-205-5p axis regulates MPP+-induced cell apoptosis in MN9D cells by directly targeting LRRK2. American journal of translational research, 10(2), 563–572.

33. Lu, Y., Gong, Z., Jin, X., Zhao, P., Zhang, Y., & Wang, Z. (2020). LncRNA MALAT1 targeting miR-124-3p regulates DAPK1 expression contributes to cell apoptosis in Parkinson's Disease. Journal of cellular biochemistry, 121(12), 4838-4848.

34. Yang, H. (2021). LncRNA MALAT1 potentiates inflammation disorder in Parkinson's disease. International Journal of Immunogenetics, 48(5), 419-428.

35. Buratti, E., and Baralle, F. E. (2001) Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J. Biol. Chem. 276, 36337-36343

36. Cohen, T. J., Lee, V. M., & Trojanowski, J. Q. (2011). TDP-43 functions and pathogenic mechanisms implicated in TDP-43 proteinopathies. Trends in molecular medicine, 17(11), 659-667.

37. Volkening, K., Leystra-Lantz, C., Yang, W., Jaffee, H., and Strong, M. J. (2009) Tar DNA binding protein of 43 kDa (TDP-43), 14-3-3 proteins and copper/zinc superoxide dismutase (SOD1) interact to modulate NFL mRNA stability. Implications for altered RNA processing in amyotrophic lateral sclerosis (ALS). *Brain Res.* **1305**, 168–182

38. Mann, J.R., Gleixner, A.M., Mauna, J.C., Gomes, E., DeChellis-Marks, M.R., Needham, P.G., Copley, K.E., Hurtle, B., Portz, B., Pyles, N.J. and Guo, L., 2019. RNA binding antagonizes neurotoxic phase transitions of TDP-43. Neuron, 102(2), pp.321-338.

39. Loganathan, S., Lehmkuhl, E. M., Eck, R. J., and Zarnescu, D. C. (2020) To Be or Not To Be…Toxic—Is RNA Association With TDP-43 Complexes Deleterious or Protective in Neurodegeneration? *Front. Mol. Biosci.* **8,** 154

40. Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Župunski, V. and Patani, R., 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nature neuroscience, 14(4), pp.452-458.

41. Lukavsky PJ, Daujotyte D, Tollervey JR, Ule J, Stuani C, Buratti E, Baralle FE, Damberger FF, Allain FH. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. Nat Struct Mol Biol. 2013 Dec;20(12):1443-9. doi: 10.1038/nsmb.2698. Epub 2013 Nov 17. PMID: 24240615.

42. Rengifo-Gonzalez, J. C., El Hage, K., Clément, M. J., Steiner, E., Joshi, V., Craveur, P., ... & Bouhss, A. (2021). The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. Elife, 10, e67605.

43. West, Jason A., Christopher P. Davis, Hongjae Sunwoo, Matthew D. Simon, Ruslan I. Sadreyev, Peggy I. Wang, Michael Y. Tolstorukov, and Robert E. Kingston. "The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites." Molecular cell 55, no. 5 (2014): 791-802.

44. Scherer, Marian, Michal Levin, Falk Butter, and Marion Scheibe. "Quantitative proteomics to identify nuclear RNA-binding proteins of Malat1." International journal of molecular sciences 21, no. 3 (2020): 1166.

45. Gaweda-Walerych, K., Sitek, E. J., Narożańska, E., & Buratti, E. (2021). Parkin beyond Parkinson's Disease—A Functional Meaning of Parkin Downregulation in TDP-43 Proteinopathies. Cells, 10(12), 3389.

46. Yamashita, Rika, Goichi Beck, Yuki Yonenobu, Kimiko Inoue, Akihiko Mitsutake, Hiroyuki Ishiura, Masato Hasegawa, Shigeo Murayama, and Hideki Mochizuki. "TDP-43 proteinopathy presenting with typical symptoms of Parkinson's disease." Mov Disord 37, no. 7 (2022): 1561-1563.

47. Gao, Ju, Luwen Wang, Mikayla L. Huntley, George Perry, and Xinglong Wang. "Pathomechanisms of TDP-43 in neurodegeneration." Journal of neurochemistry 146, no. 1 (2018): 7-20.

48. Panda, A. C., Martindale, J. L., & Gorospe, M. (2017). Polysome Fractionation to Analyze mRNA Distribution Profiles. Bio-protocol, 7(3), e2126. https://doi.org/10.21769/BioProtoc.2126

49. Hallegger, M., Chakrabarti, A.M., Lee, F.C., Lee, B.L., Amalietti, A.G., Odeh, H.M., Copley, K.E., Rubien, J.D., Portz, B., Kuret, K. and Huppertz, I., 2021. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell, 184(18), pp.4680-4696.

50. Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E. and Baralle, F.E., 2011. TDP-43 regulates its mRNA levels through a negative feedback loop. The EMBO journal, 30(2), pp.277-288.

51. Mann, J. R., Gleixner, A. M., Mauna, J. C., Gomes, E., DeChellis-Marks, M. R., Needham, P. G., ... & Donnelly, C. J. (2019). RNA binding antagonizes neurotoxic phase transitions of TDP-43. Neuron, 102(2), 321-338.

52. Mann, M., Wright, P. R., & Backofen, R. (2017). IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. Nucleic acids research, 45(W1), W435–W439. https://doi.org/10.1093/nar/gkx279

53. Zuker, M. and Stiegler, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information Nucleic Acid Res. 9(1): 133-148, 1981

54. Mückstein, Ulrike, Hakim Tafer, Jörg Hackermüller, Stephan H. Bernhart, Peter F. Stadler, and Ivo L. Hofacker. "Thermodynamics of RNA–RNA binding." Bioinformatics 22, no. 10 (2006): 1177-1182.

55. Button, Aileen C., Simone D. Hall, Ethan L. Ashley, and Colleen A. McHugh. "Dissection of protein and RNA regions required for SPEN binding to XIST A-repeat RNA." RNA 30, no. 3 (2024): 240-255.

56. Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P., & Herschlag, D. (2020). How to measure and evaluate binding affinities. eLife, 9, e57264. https://doi.org/10.7554/eLife.57264

57. Ryder, S. P., Recht, M. I., & Williamson, J. R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. Methods in molecular biology (Clifton, N.J.), 488, 99–115. https://doi.org/10.1007/978-1-60327-475-3_7

58. Kalivendi, S. V., Kotamraju, S., Cunningham, S., Shang, T., Hillard, C. J., and Kalyanaraman, B. (2003) 1-Methyl-4-phenylpyridinium (MPP+)-induced apoptosis and mitochondrial oxidant generation: role of transferrin-receptor-dependent iron and hydrogen peroxide. Biochem. J. 371, 151-164

59. Cassarino, D. S., Parks, J. K., Parker, W. D. Jr., and Bennett, J. P. Jr. (1999) The parkinsonian neurotoxin MPP+ opens the mitochondrial permeability transition pore and releases cytochrome c in isolated mitochondria via an oxidative mechanism. Biochim. Biophys. Acta. 1453, 49-62

60. Lu, Y., Gong, Z., Jin, X., Zhao, P., Zhang, Y., and Wang, Z. (2020) LncRNA MALAT1 targeting miR-124-3p regulates DAPK1 expression contributes to cell apoptosis in Parkinson's Disease. J. Cell. Biochem. 121, 4838-4848

61. Martin G, Gruber AR, Keller W, Zavolan M. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. Cell Rep. 2012 Jun 28;1(6):753-63. doi: 10.1016/j.celrep.2012.05.003. Epub 2012 Jun 7. PMID: 22813749.

62. Movassat, M., Crabb, T. L., Busch, A., Yao, C., Reynolds, D. J., Shi, Y., & Hertel, K. J. (2016). Coupling between alternative polyadenylation and alternative splicing is limited to terminal introns. RNA biology, 13(7), 646–655. https://doi.org/10.1080/15476286.2016.1191727

63. Grozdanov, Petar N., Elahe Masoumzadeh, Vera M. Kalscheuer, Thierry Bienvenu, Pierre Billuart, Marie-Ange Delrue, Michael P. Latham, and Clinton C. MacDonald. "A missense mutation in the CSTF2 gene that impairs the function of the RNA recognition motif and causes defects in 3′ end processing is associated with intellectual disability in humans." Nucleic acids research 48, no. 17 (2020): 9804-9821.

64. Paul MS, Michener SL, Pan H, Chan H, Pfliger JM, Rosenfeld JA, Lerma VC, Tran A, Longley MA, Lewis RA, Weisz-Hubshman M, Bekheirnia MR, Bekheirnia N, Massingham L, Zech M, Wagner M, Engels H, Cremer K, Mangold E, Peters S, Trautmann J, Mester JL, Guillen Sacoto MJ, Person R, McDonnell PP, Cohen SR, Lusk L, Cohen ASA, Le Pichon JB, Pastinen T, Zhou D, Engleman K, Racine C, Faivre L, Moutton S, Denommé-Pichon AS, Koh HY, Poduri A, Bolton J, Knopp C, Julia Suh DS, Maier A, Toosi MB, Karimiani EG, Maroofian R, Schaefer GB, Ramakumaran V, Vasudevan P, Prasad C, Osmond M, Schuhmann S, Vasileiou G, Russ-Hall S, Scheffer IE, Carvill GL, Mefford H; Undiagnosed Diseases Network; Bacino CA, Lee BH, Chao HT. A syndromic neurodevelopmental disorder caused by rare variants in PPFIA3. Am J Hum Genet. 2024 Jan 4;111(1):96-118. doi: 10.1016/j.ajhg.2023.12.004. Erratum in: Am J Hum Genet. 2024 Apr 4;111(4):805. PMID: 38181735; PMCID: PMC10806447.

65. Wong MY, Liu C, Wang SSH, Roquas ACF, Fowler SC, Kaeser PS. Liprin-α3 controls vesicle docking and exocytosis at the active zone of hippocampal synapses. Proc Natl Acad Sci U S A. 2018 Feb 27;115(9):2234-2239. doi: 10.1073/pnas.1719012115. Epub 2018 Feb 8. PMID: 29439199; PMCID: PMC5834710.

66. Zhen M, Jin Y. The liprin protein SYD-2 regulates the differentiation of presynaptic termini in C. elegans. *Nature.* 1999 Sep 23;401(6751):371-5. doi: 10.1038/43886. PMID: 10517634.

67. Joshi CS, Suryawanshi AR, Khan SA, Balasinor NH, Khole VV. Liprin α3: a putative estrogen regulated acrosomal protein. Histochem Cell Biol. 2013 Apr;139(4):535-48. doi: 10.1007/s00418-012-1044-y. Epub 2012 Nov 3. PMID: 23124857.

68. Serra-Pagès C, Medley QG, Tang M, Hart A, Streuli M. Liprins, a family of LAR transmembrane protein-tyrosine phosphatase-interacting proteins. J Biol Chem. 1998 Jun 19;273(25):15611-20. doi: 10.1074/jbc.273.25.15611. PMID: 9624153.

69. Hemphill, Wayne O., Calvin K. Voong, Regan Fenske, James A. Goodrich, and Thomas R. Cech. "Multiple RNA-and DNA-binding proteins exhibit direct transfer of polynucleotides with implications for target-site search." Proceedings of the National Academy of Sciences 120, no. 26 (2023): e2220537120.

70. Lightfoot HL, Hall J. Endogenous polyamine function--the RNA perspective. Nucleic Acids Res. 2014;42(18):11275-11290. doi:10.1093/nar/gku837

71. Marton, Laurence J., and Anthony E. Pegg. "Polyamines as targets for therapeutic intervention." Annual review of pharmacology and toxicology 35, no. 1 (1995): 55-91.

72. Bachmann, André S., and Victor A. Levin. "Clinical applications of polyamine-based therapeutics." (2011).

73. Casero Jr, Robert A., Tracy Murray Stewart, and Anthony E. Pegg. "Polyamine metabolism and cancer: treatments, challenges and opportunities." Nature Reviews Cancer 18, no. 11 (2018): 681-695.

74. Smith, Tom, Andreas Heger, and Ian Sudbery. "UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy." Genome research 27, no. 3 (2017): 491-499.

75. Martin, Marcel. "Cutadapt removes adapter sequences from high-throughput sequencing reads." EMBnet. journal 17, no. 1 (2011): 10-12.

76. Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. "STAR: ultrafast universal RNA-seq aligner." Bioinformatics 29, no. 1 (2013): 15-21.

77. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. and Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience, 10(2), p.giab008.

78. Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. "Integrative genomics viewer." Nature biotechnology 29, no. 1 (2011): 24-26.

79. Gagnon, Keith T., Liande Li, Bethany A. Janowski, and David R. Corey. "Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading." Nature protocols 9, no. 9 (2014): 2045-2060.

80. Gagliardi, Miriam, and Maria R. Matarazzo. "Rip: Rna Immunoprecipitation." Polycomb Group Proteins: Methods and Protocols (2016): 73-86.

81. Raran-Kurussi, Sreejith, Scott Cherry, Di Zhang, and David S. Waugh. "Removal of affinity tags with TEV protease." Heterologous Gene Expression in E. coli: Methods and Protocols (2017): 221-230.

82. [pre-print] Su, T., Trang, N., Zhu, J., Kong, L., Cheung, D., Chou, V., Ellis, L., Huang, C., Camden, N. and McHugh, C.A., 2023. GRAS1 non-coding RNA protects against DNA damage and cell death by binding and stabilizing NKAP. bioRxiv, pp.2023-06.

Acknowledgements

We appreciate the continued support of our funding agencies and the Department of Chemistry and Biochemistry at UCSD. The authors acknowledge Biorender.com for the creation of visual figures used in the study.

This work was supported by NIH grant R00 GM120494 to C.A.M., the San Diego Fellowship to A.B., UCSD Chemistry and Biochemistry faculty startup funds to C.A.M, and the UCSD Dean of Physical Sciences Summer Undergraduate Research Award to R.J.

Chapter 3, in full, is currently in preparation for submission for publication of the material. This material includes author contributions from Raul Johnson, Jonathan Zhu, Lauren Ellis, Simone Hall, and Colleen A. McHugh. The dissertation author was the primary researcher and author of this material.

Figures

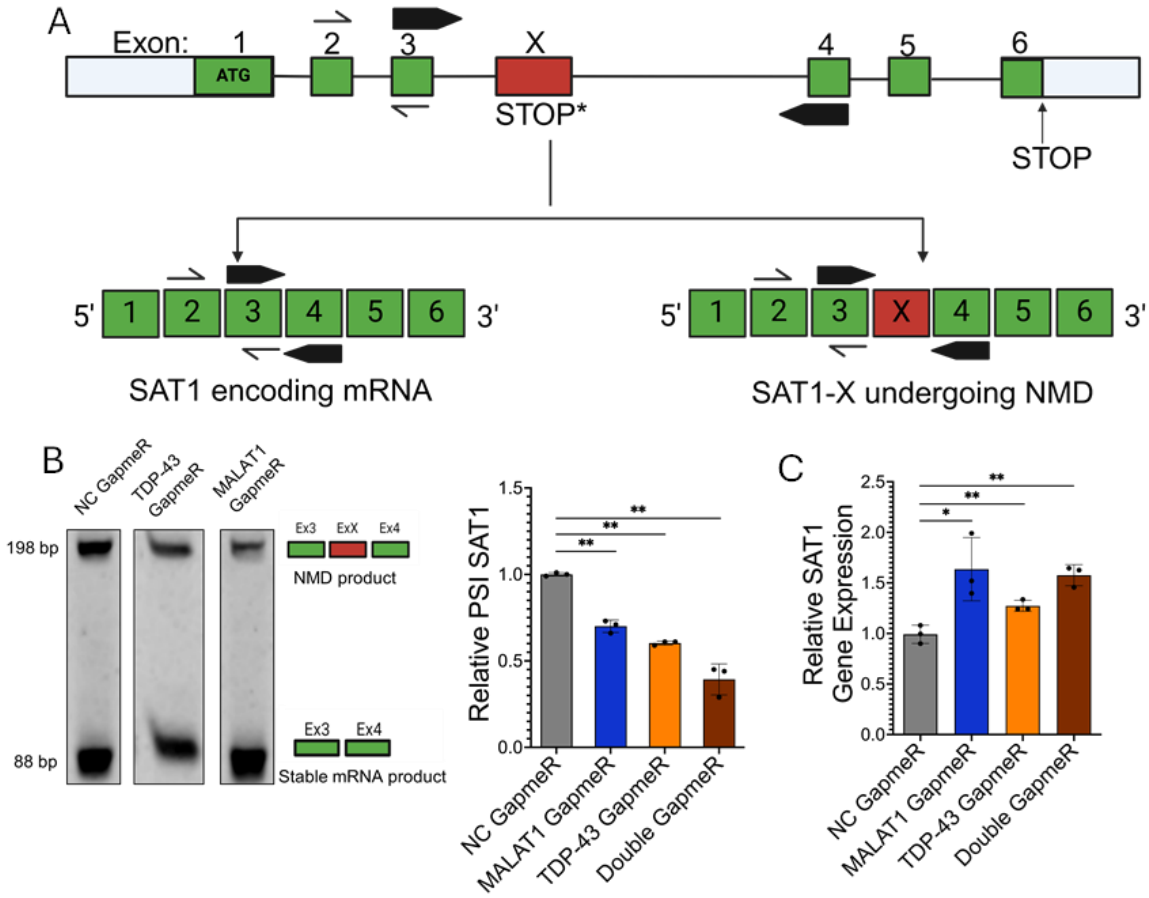


Figure 3.1 – MALAT1 and TDP-43 regulate alternative splicing of SAT1 pre-mRNA

A) Overview of SAT1 alternative splicing mechanism. Exons are depicted in green, and the alternatively spliced NMD-inducing exon 'X' is noted in red. PCR primers are shown with full arrows and qPCR primers are shown with half arrows. B) Splicing gel and relative percent splice inclusion (PSI) analysis of SAT1 in knockdown conditions and visual representation. N=3 C) SAT1 RNA expression levels in knockdown conditions, relative to GAPDH RNA levels. All experiments for figure 1 are conducted in HEK293 cells. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

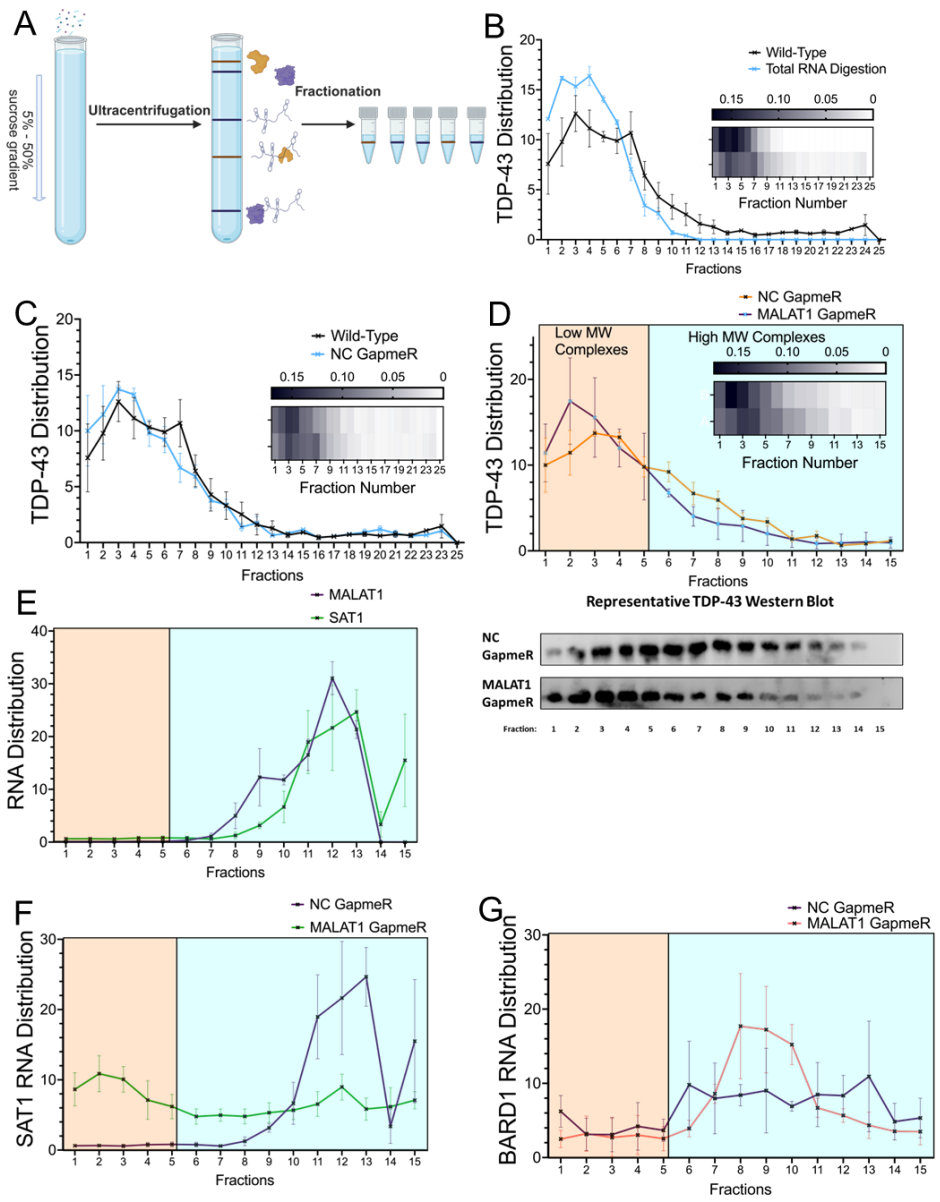


Figure 3.2 - TDP-43 and SAT1 mRNA form high molecular weight complexes that are MALAT1 dependent in cells

(A) Schematic overview of sucrose gradient ultracentrifugation (B) TDP-43 protein distribution in sucrose gradient between wild-type and total RNA digestion. Protein localization has also been plotted as a heat map. N=3. (C) Wild-type and NC GapmeR conditions generate similar TDP-43 stable complexes in sucrose gradient. Protein localization has also been plotted as a heat map. N=3. (D) MALAT1 knockdown relocalizes TDP-43 to lower dense regions of the gradient, forming more lower MW complexes. Protein localization has also been plotted as a heat map. A representative western blot for TDP-43 is shown for collected fractions in NC and MALAT1 GapmeR conditions. N=3. (E) RNA distribution of MALAT1 and SAT1 mRNA in sucrose gradient under NC GapmeR conditions. Both RNAs localize in similar densities in the gradient. N=3. (F) RNA distribution of SAT1 mRNA in sucrose gradient under NC and MALAT1 GapmeR conditions, showing significant SAT1 mRNA re-localization in MALAT1 GapmeR conditions. N=3. (G) RNA localization of BARD1 mRNA in sucrose gradient under NC and MALAT1 GapmeR conditions, showing no significant re-localization. All experiments in figure 2 are conducted in K562 cells. data are plotted with error bars representing SD.

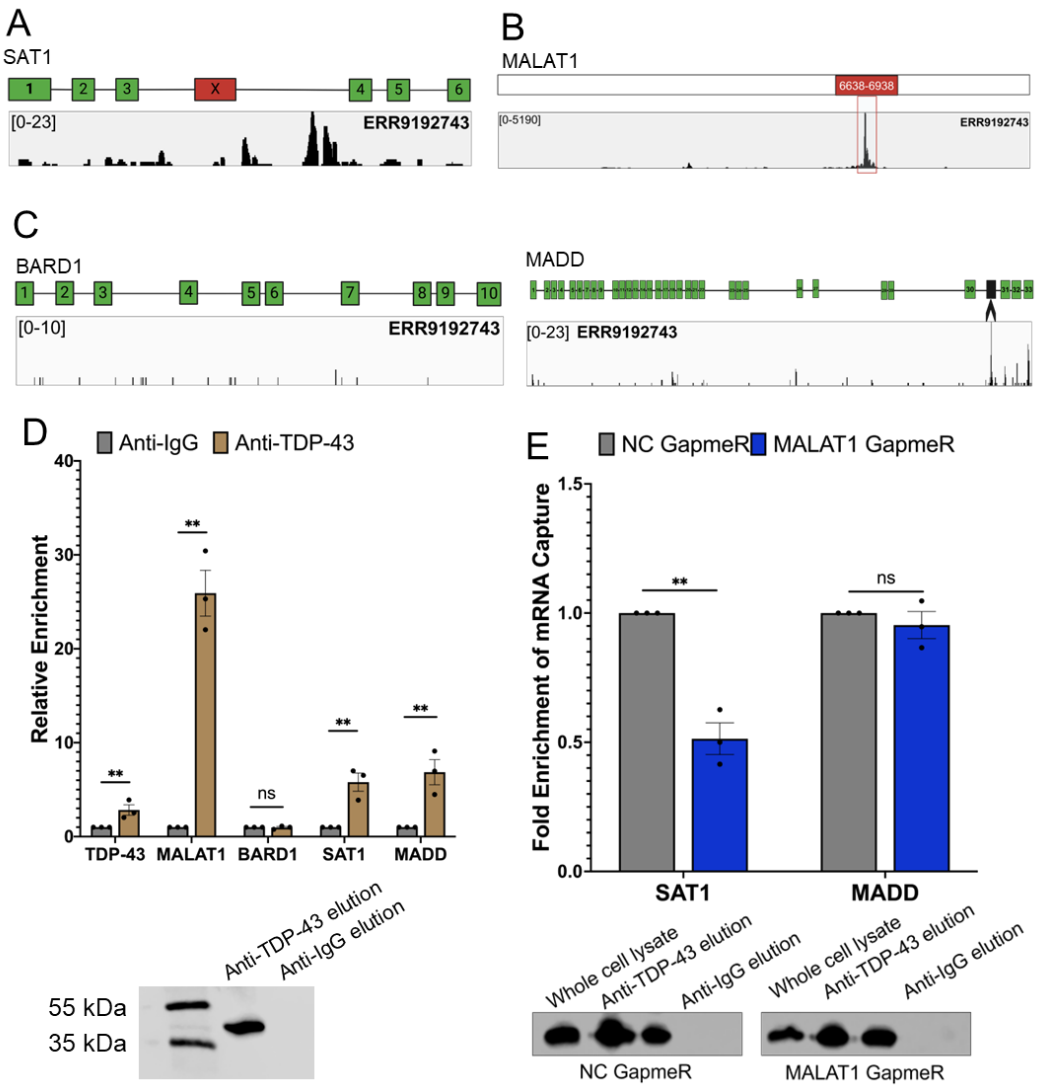


Figure 3.3 - TDP-43 binds directly to SAT1 mRNA and MALAT1 RNA transcripts, and loss of MALAT1 reduces binding of SAT1 mRNA to TDP-43

(A) TDP-43 CLIP-seq identification of direct binding site on SAT1 mRNA in HEK293 Cells. There is a strong binding peak on intron 4 (in4). (B) TDP-43 CLIP-seq identification of direct binding site on MALAT1 in HEK293 Cells. There is a strong 300nt binding peak encompassing nucleotides 6638-6938. (C) TDP-43 CLIP-seq identification of direct binding site on negative control BARD1 mRNA and positive control MADD mRNA in HEK293 Cells. There is no visible direct binding of TDP-43 to BARD1 mRNA, and there is a strong binding peak on intron 30 (in30) of MADD mRNA. (D) Quantitative RNA-IP on HEK293 cells with Anti-IgG and Anti-TDP-43 antibodies show TDP-43 binding enrichment to target transcripts. A representative western blot is shown to validate TDP-43 protein capture. (E) Quantitative RNA-IP on HEK293 cells in NC and MALAT1 GapmeR conditions reveal a loss in TDP-43 binding to SAT1 mRNA, and no significant change in binding to MADD mRNA. Representative western blots are shown to validate equivalent TDP-43 protein capture in both conditions. N =3. All experiments in figure 3 are performed in HEK293 cell lines. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

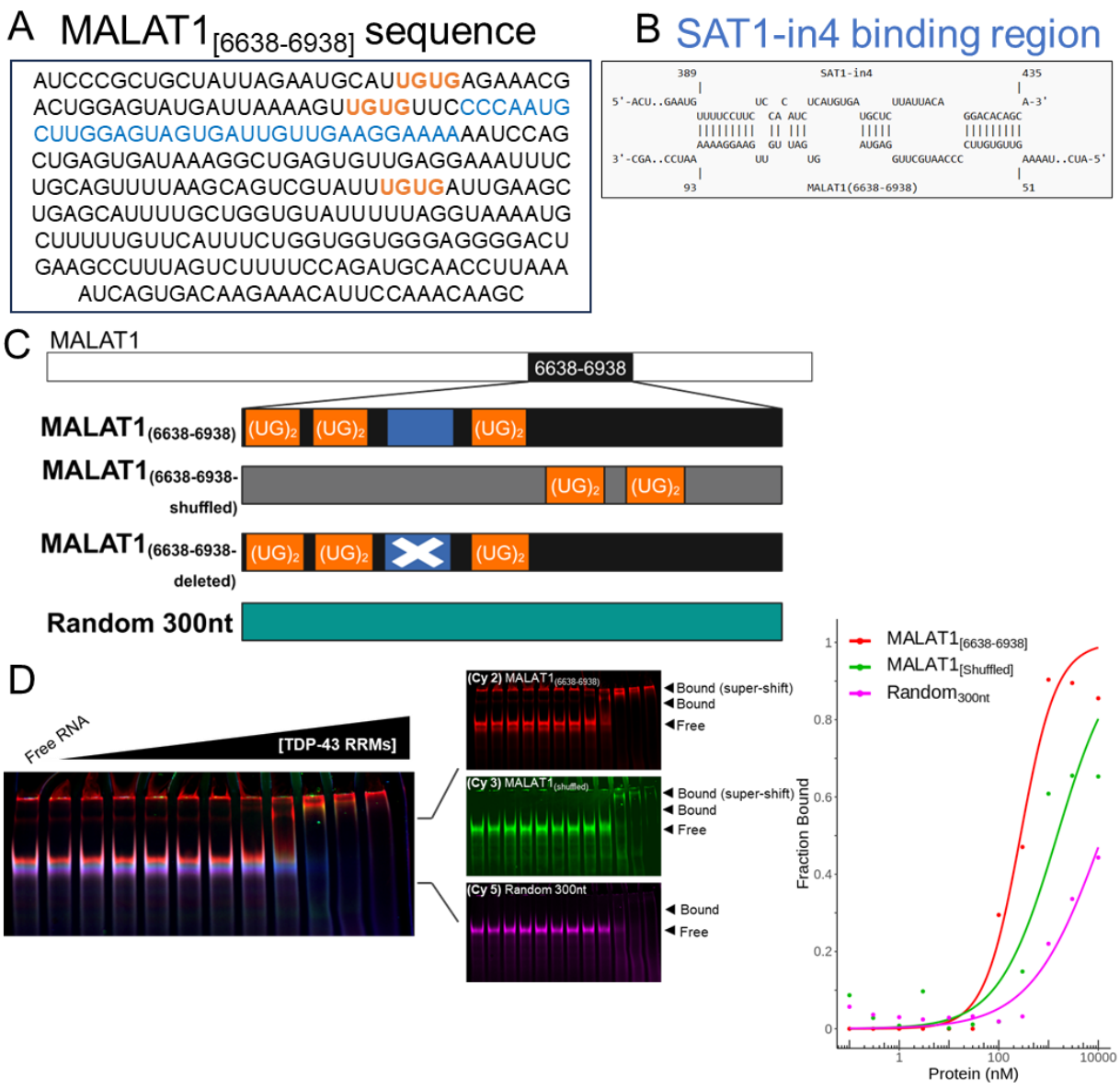


Figure 3.4 - MALAT1 non-coding RNA has specific binding sites for TDP-43 protein and SAT1 pre-mRNA

(A) RNA sequence identity of MALAT1$_{[6639\text{-}6938]}$. Tandem UG repeats are highlighted in orange. Predicted RNA-RNA hybridization sequence between MALAT1 and SAT1-in4 mRNA is highlighted in blue. (B) Computational prediction of watson-crick based MALAT1-SAT1in4 hybridization generated by the Inta-RNA program. (C) Overview of the four MALAT1$_{[6638\text{-}6938]}$-based construct designs used for multiplexed EMSA assays. (D) Multiplexed EMSA of fluorescently-labelled MALAT1$_{[6638\text{-}6938]}$, MALAT1$_{[6638\text{-}6938\text{-}shuffled]}$ and a randomly generated 300nt sequence with TDP-43 RRMS. Visualization and quantification of Kd is provided in the graph and Table 1.

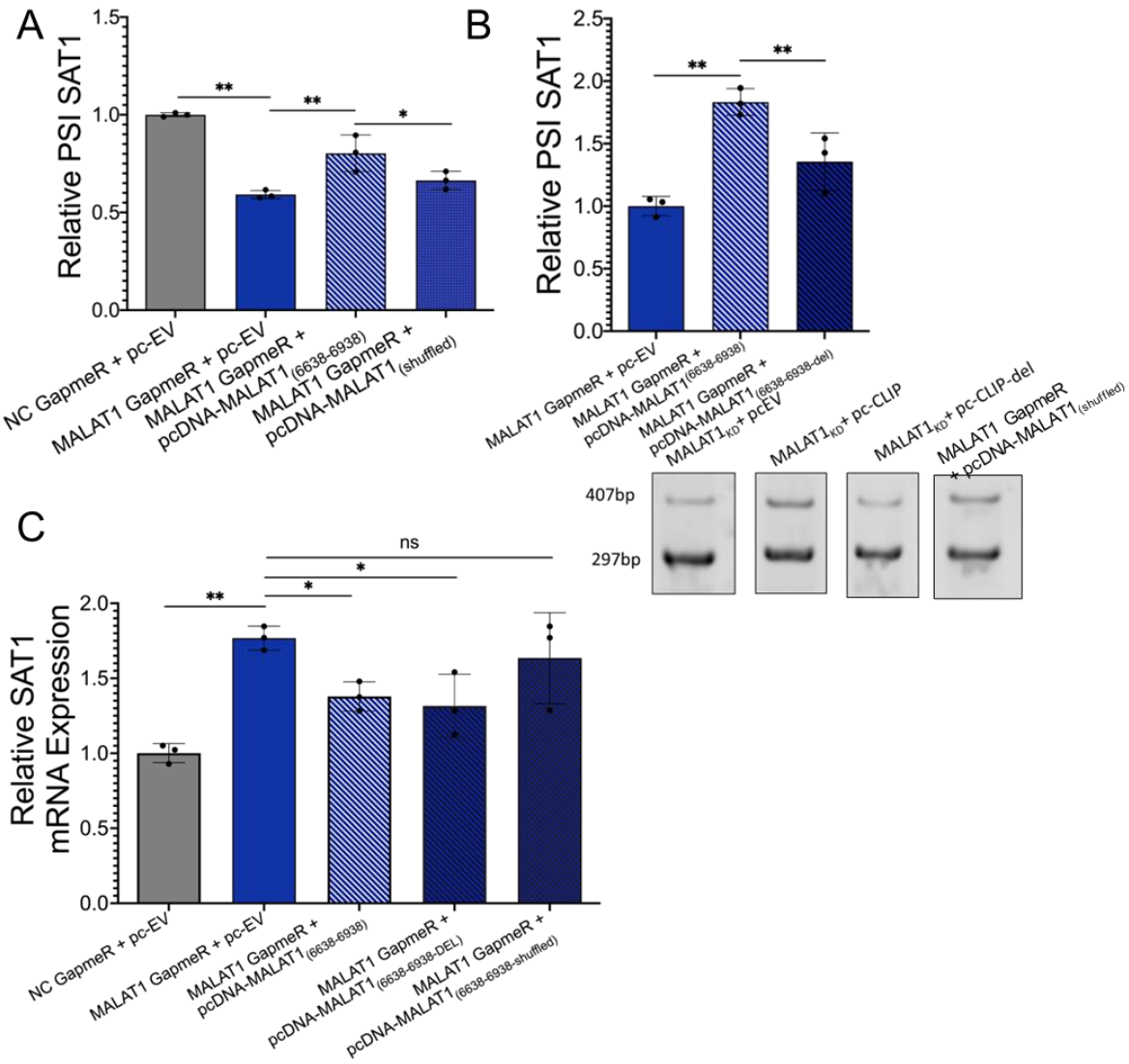


Figure 3.5 - MALAT1$_{[6638-6938]}$ binding to SAT1 pre-mRNA mediates efficient SAT1 alternative splicing

(A) Relative PSI of SAT1 mRNA in MALAT1 GapmeR conditions supplemented with ectopic overexpression of pc-EV, pc-MALAT1$_{[6638-6938]}$ and pc-MALAT1$_{[6638-6938-shuffled]}$. N=3. (B) Splicing gel and PSI analysis of SAT1 mRNA in MALAT1 GapmeR conditions supplemented with ectopic overexpression of pc-MALAT1$_{[6638-6938]}$ and pc-MALAT1$_{[6638-6938-Del]}$. Representative splicing gel lanes are shown. N=3. (C) Gene expression analysis of SAT1 in MALAT1 GapmeR conditions supplemented with ectopic overexpression of pc-MALAT1 vectors. N=3. All experiments for figure 5 are conducted in HEK293 cells. All Statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

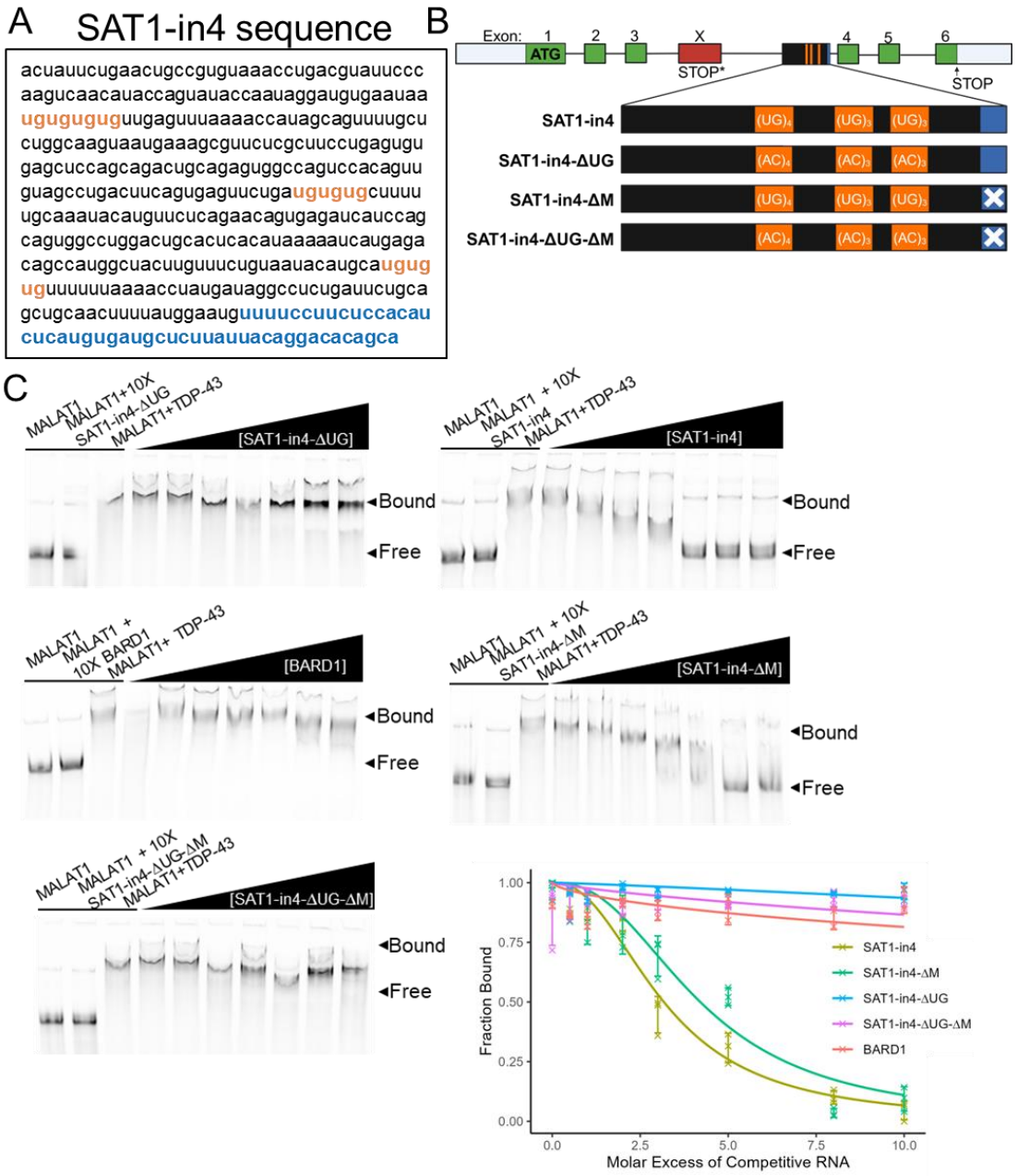


Figure 3.6 - SAT1 premRNA binding to MALAT1 and TDP-43 occurs through sequence-specific interactions

(A) RNA sequence identity of SAT1-in4 mRNA. Tandem UG repeats are highlighted in orange. Predicted RNA-RNA hybridization sequence between MALAT1 and SAT1-in4 mRNA is highlighted in blue. (B) Overview of the four SAT1-in4 based constructs used for in-vitro competitive EMSA experiments. (C) SAT1-in4-ΔUG, SAT1-in4-ΔUG-ΔM and BARD1 cannot outcompete TDP-43-MALAT1$_{[6638-6938]}$ binding due to a lack of UG repeats. SAT1-in4 outcompetes TDP-43-MALAT1$_{[6638-6938]}$ binding faster than SAT1-in4-ΔM. Quantification of competitive EMSA data is shown through graphical representation. N=3.

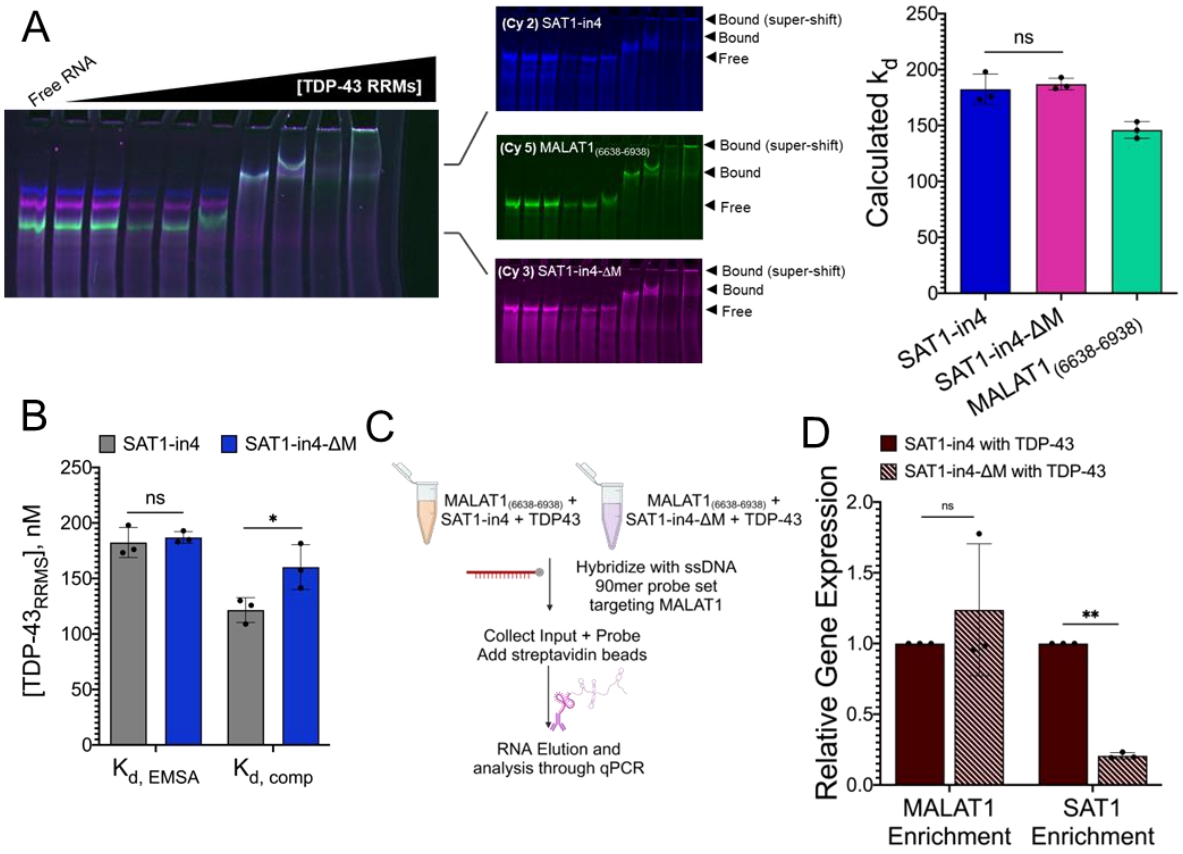


Figure 3.7 - MALAT1 enhances binding of TDP-43 to SAT1 pre-mRNA

(A) Multiplexed EMSA of fluorescently-labelled SAT1-in4, SAT1-in4-ΔM, and MALAT1$_{[6638-6938]}$ RNA with TDP-43 RRMs indicates that SAT1-in4 and SAT1-in4-ΔM both have comparable k_d values. N=3. (B) A significant difference is observed in binding affinity of SAT1-in4 compared to SAT1-in4-ΔM when k_d is calculated in competitive EMSA conditions ($k_{d,\,comp}$) using the Lin and Riggs equation. N=3. (C) Schematic overview of in-vitro assembly and pulldown of MALAT1-TDP-43-SAT1mRNA complexes using MALAT1 ssDNA probes. (D) Pulldown of in-vitro assembled MALAT1-TDP-43-SAT1mRNA complexes indicate a loss in SAT1mRNA-MALAT1 binding between SAT1-in4 and SAT1-in4-ΔM constructs. N=3. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

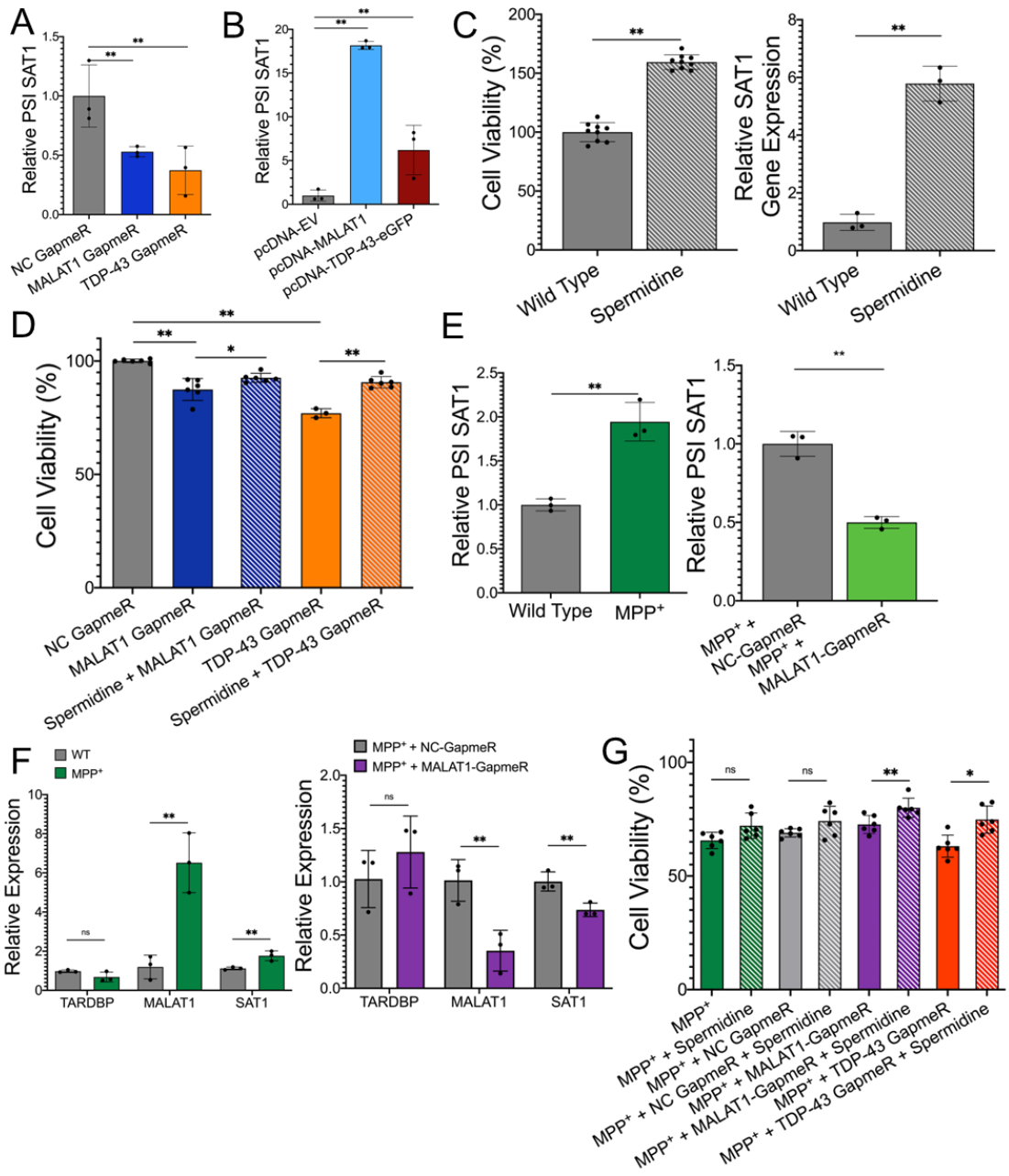


Figure 3.8 - MALAT1 and TARDBP RNA knockdown through GapmeR treatment affect survival of SH-SY5Y neuroblastoma cells

(A) Relative PSI of SAT1 mRNA in MALAT1 and TDP-43 GapmeR conditions. N=3. (B) Relative PSI of SAT1 mRNA in MALAT1 and TDP-43 ectopic overexpression conditions. N=3 (C) Cell viability and SAT1 mRNA expression in SH-SY5Y cells with and without spermidine addition. N=3. (D) Cell viability of MALAT1 and TDP-43 GapmeR conditions in SH-SY5Y cells with and without spermidine addition. N=6. (E) Relative PSI of SAT1 mRNA increases in MPP^+ conditions and decreases with addition of MALAT1 GapmeR compared to NC GapmeR. N=3. (F) Gene expression of MALAT1 and SAT1 mRNA increases in MPP^+ conditions and decreases with addition of MALAT1 GapmeR compared to NC GapmeR. N=3. (G) Cell viability of MALAT1 and TDP-43 GapmeR conditions in MPP^+ conditions with and without spermidine addition. N=6. All experiments for figure 8 are conducted in SH-SY5Y cells. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

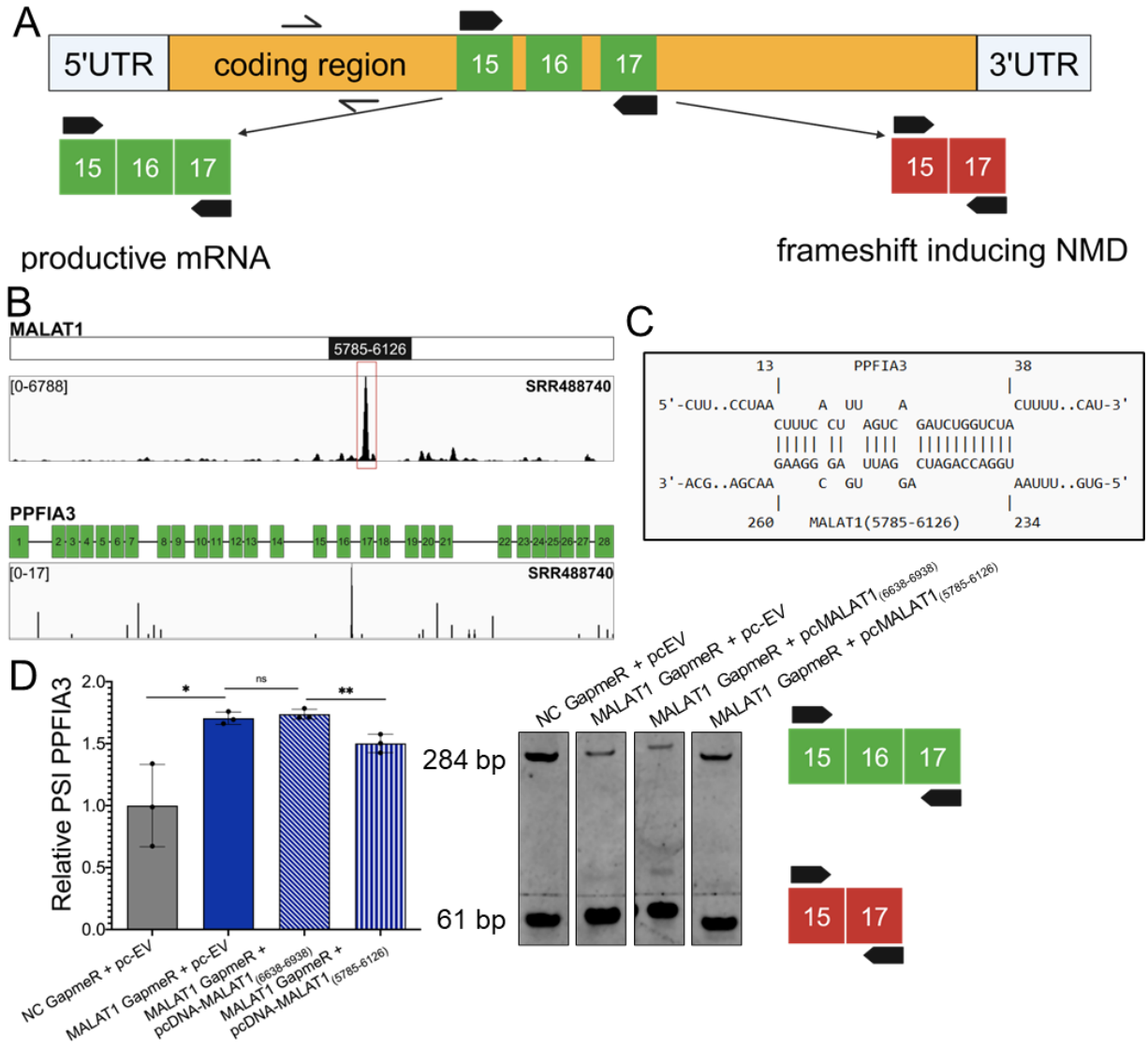


Figure 3.9 - MALAT1 similarly enhances alternative splicing of liprin-α (PPFIA3) by CSTF2 through sequence-specific interactions

(A) Overview of PPFIA3 alternative splicing mechanism. PCR primers are shown with full arrows. Exons and the productive mRNA are denoted in green. The NMD-inducing mRNA is denoted in red. (B) CLIP-seq identification of CSTF2 direct binding site in HEK293 cells on MALAT1$_{[5785-6126]}$ and on PPFIA3-in16 pre-mRNA. (C) Computational prediction of watson-crick based MALAT1-PPFIA3-in16 hybridization generated by the Inta-RNA program. (D) Relative PSI of PPFIA3 mRNA in MALAT1 GapmeR conditions with ectopic overexpression of pc-MALAT1$_{[6638-6938]}$ or pc-MALAT1$_{[5785-6126]}$. A representative splicing gel is also shown. N=3. All experiments for Figure 9 are conducted in HEK293 cells. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.

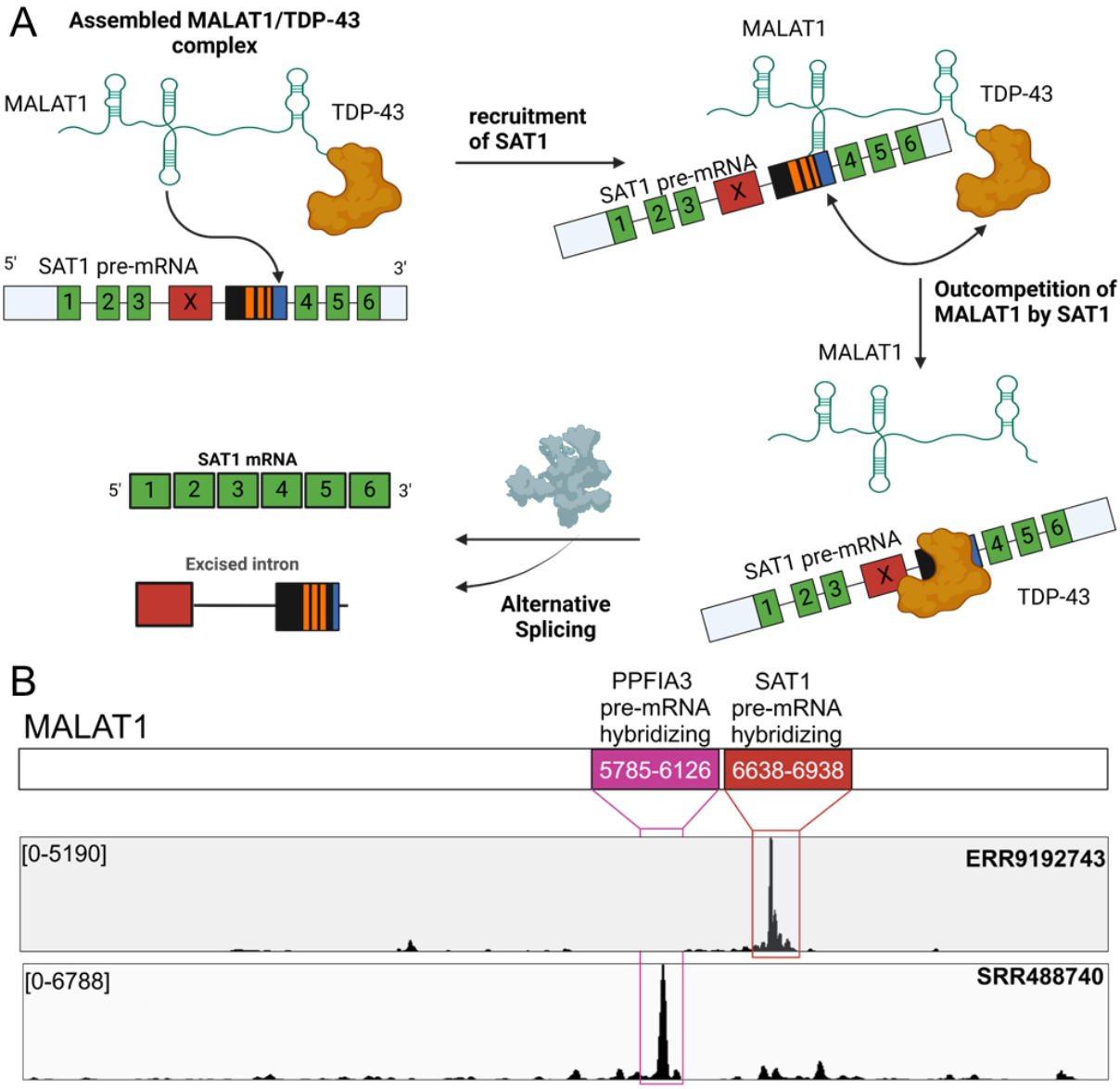


Figure 3.10 - Proposed model of MALAT1 regulation of TDP-43 alternative splicing activity

(A) Schematic overview of proposed model through which MALAT1 regulates TDP-43 binding to intronic regions of target pre-mRNA for AS. (B) TDP-43 CLIP-seq overview of direct binding on MALAT1 showing distinct binding locations for AS regulation of SAT1 pre-mRNA and PPFIA3 pre-mRNA.

Supplemental Tables

Supplemental Table 3.1 - Nucleotide sequences of MALAT1$_{[6638-6938]}$-based constructs

Construct	RNA Sequence
MALAT1 **[6638-6938]**	AUCCCGCUGCUAUUAGAAUGCAUUGUGAAACGACUGGAGUAUGAUU AAAAGUUGUGUUCCCCAAUGCUUGGAGUAGUGAUUGUUGAAGGAA AAAAUCCAGCUGAGUGAUAAAGGCUGAGUGUUGAGGAAAUUUCUG CAGUUUUAAGCAGUCGUAUUUGUGAUUGAAGCUGAGCAUUUUGCU GGUGUAUUUUUAGGUAAAAUGCUUUUUGUUCAUUUCUGGUGGUGG GAGGGGACUGAAGCCUUUAGUCUUUUCCAGAUGCAACCUUAAAAUC AGUGACAAGAAACAUUCCAAACAAGC
MALAT1 **[6638-6938-Shuffled]**	GGAACAAGAUCUGUUUGUAUUAAUUUUUAUGUUGGAAAGUUUUUC UAAUGACCGUUGAGUACUAUGAAGGGUAGUUAAAAUACGACCGGC GUAGGCGUAAAAUACAAGAUGGCUGCUAUUCAAUUGGCGUUACGGC CAAUACAUAGUGGUGGUUCAGAUACUAGUUCUUUAAUCAAAGGUA AUUGUUUAGAAGCCGACCGGUUUUGUGCGAUCUUUGUGAGAUGAG AUUAGCACUUUUCUCCUGCUGAUACUUAGGUUGAUAUUGAAAAAG AAUAUAGGGUCGCGUCCGAAGACCAUGGG
MALAT1 **[6638-6938-del]**	AUCCCGCUGCUAUUAGAAUGCAUUGUGAAACGACUGGAGUAUGAUU AAAAAAUCCAGCUGAGUGAUAAAGGCUGAGUGUUGAGGAAAUUUC UGCAGUUUUAAGCAGUCGUAUUUGUGAUUGAAGCUGAGUACAUUU UGCUGGUGUAUUUUUAGGUAAAAUGCUUUUUGUUCAUUUCUGGUG GUGGGAGGGGACUGAAGCCUUUAGUCUUUUCCAGAUGCAACCUUAA AAUCAGUGACAAGAAACAUUCCAAACAAGC
Random 300nt	GGUAGCAGAUGCGUCUUAAAUAAUACUCAAGUGGUAUCUUUACCAU UUCCGCCAAUCUAACGGAGGCUCAUGUCCGACACGUGAUAAUCGCU GACCGACUAAUAGCGUCCAGCCGAUGGGCUACUGUCGCUAUAAGCG UCCGCCCGCGCUUAAUUAUGACGACGCGGCUCUCACCACUCAGAUC UUGGUACUCCCCCCAAUCAAACGGUGAGAUUCCUUGAAUGAGACGC CAUCACUCGGAUAGCCUGCCCUCAUCUGAUAAGCGCGAUAAUAUGC GCAGACAAGUCAAACCCGGUCAGC

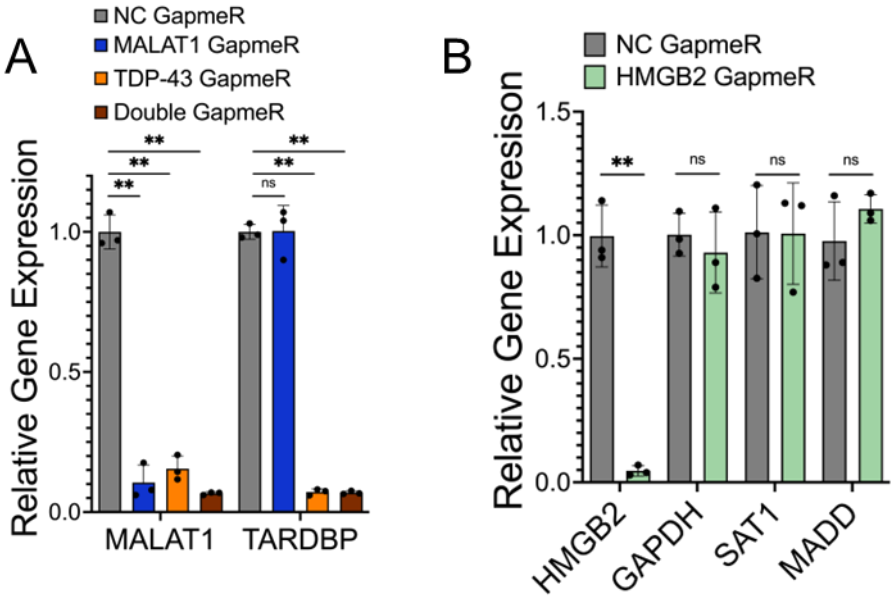
Supplemental Table 3.2 - $K_{d, EMSA}$ for $MALAT1_{[6638-6938]}$ fragmented RNA constructs

Fragment identity	**$K_{d, EMSA}$**
$MALAT1_{[6638-6738]}$	408.1
$MALAT1_{[6678-6778]}$	533.2
$MALAT1_{[6763-6863]}$	466.1

Supplemental Table 3.3 - Nucleotide sequences of SAT1-in4-based constructs

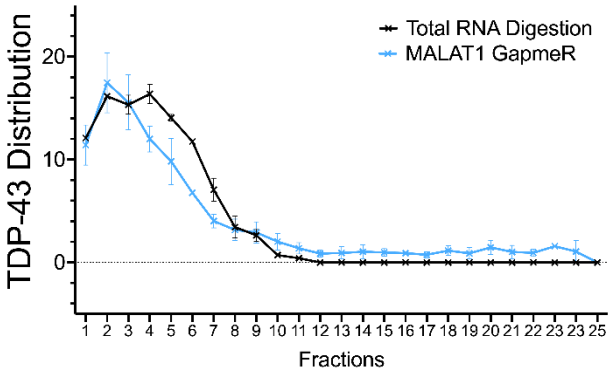
Construct	**RNA Sequence**
SAT1-in4	ACUAUUCUGAACUGCCGUGUAAACCUGACGUAUUCCCAAGUCAACAUACCAGUAU ACCAAUAGGAUGUGAAUAAUGUGUGUGUUGAGUUUAAAACCAUAGCAGUUUUGCU CUGGCAAGUAAUGAAAGCGUUCUCGCUUCCUGAGUGUGAGCUCCAGCAGACUGCA GAGUGGCCAGUCCACAGUUGUAGCCUGACUUCAGUGAGUUCUGAUGUGUGCUUUU UGCAAAUACAUGUUCUCAGAACAGUGAGAUCAUCCAGCAGUGGCCUGGACUGCAC UCACAUAAAAAUCAUGAGACAGCCAUGGCUACUUGUUUCUGUAAUACAUGCAUGU GUGUUUUUUAAAACCUAUGAUAGGCCUCUGAUUCUGCAGCUGCAACUUUUAUGGA AUGUUUUCCUUCUCCACAUCUCAUGUGAUGCUCUUAUUACAGGACACAGCA
SAT1-in4-ΔM	ACUAUUCUGAACUGCCGUGUAAACCUGACGUAUUCCCAAGUCAACAUACCAGUAU ACCAAUAGGAUGUGAAUAACACACACAUUGAGUUUAAAACCAUAGCAGUUUUGCU CUGGCAAGUAAUGAAAGCGUUCUCGCUUCCUGAGUGUGAGCUCCAGCAGACUGCA GAGUGGCCAGUCCACAGUUGUAGCCUGACUUCAGUGAGUUCUGACACACACUUUU UGCAAAUACAUGUUCUCAGAACAGUGAGAUCAUCCAGCAGUGGCCUGGACUGCAC UCACAUAAAAAUCAUGAGACAGCCAUGGCUACUUGUUUCUGUAAUACAUGCACAC ACAUUUUUUAAAACCUAUGAUAGGCCUCUGAUUCUGCAGCUGCAACUUUUAUGGA AUG~~UUUUCCUUCUCCACAUCUCAUGUGAUGCUCUUAUUACAGGACACAGCA~~
SAT1-in4-ΔUG	ACUAUUCUGAACUGCCGUGUAAACCUGACGUAUUCCCAAGUCAACAUACCAGUAU ACCAAUAGGAUGUGAAUAAUGUGUGUGUUGAGUUUAAAACCAUAGCAGUUUUGCU CUGGCAAGUAAUGAAAGCGUUCUCGCUUCCUGAGUGUGAGCUCCAGCAGACUGCA GAGUGGCCAGUCCACAGUUGUAGCCUGACUUCAGUGAGUUCUGAUGUGUGCUUUU UGCAAAUACAUGUUCUCAGAACAGUGAGAUCAUCCAGCAGUGGCCUGGACUGCAC UCACAUAAAAAUCAUGAGACAGCCAUGGCUACUUGUUUCUGUAAUACAUGCAUGU GUGUUUUUUAAAACCUAUGAUAGGCCUCUGAUUCUGCAGCUGCAACUUUUAUGGA AUG~~UUUUCCUUCUCCACAUCUCAUGUGAUGCUCUUAUUACAGGACACAGCA~~
SAT1-in4-ΔUG-ΔM	ACUAUUCUGAACUGCCGUGUAAACCUGACGUAUUCCCAAGUCAACAUACCAGUAU ACCAAUAGGAUGUGAAUAACACACACAUUGAGUUUAAAACCAUAGCAGUUUUGCU CUGGCAAGUAAUGAAAGCGUUCUCGCUUCCUGAGUGUGAGCUCCAGCAGACUGCA GAGUGGCCAGUCCACAGUUGUAGCCUGACUUCAGUGAGUUCUGACACACACUUUU UGCAAAUACAUGUUCUCAGAACAGUGAGAUCAUCCAGCAGUGGCCUGGACUGCAC UCACAUAAAAAUCAUGAGACAGCCAUGGCUACUUGUUUCUGUAAUACAUGCACAC ACAUUUUUUAAAACCUAUGAUAGGCCUCUGAUUCUGCAGCUGCAACUUUUAUGGA AUG~~UUUUCCUUCUCCACAUCUCAUGUGAUGCUCUUAUUACAGGACACAGCA~~

Supplemental Figures



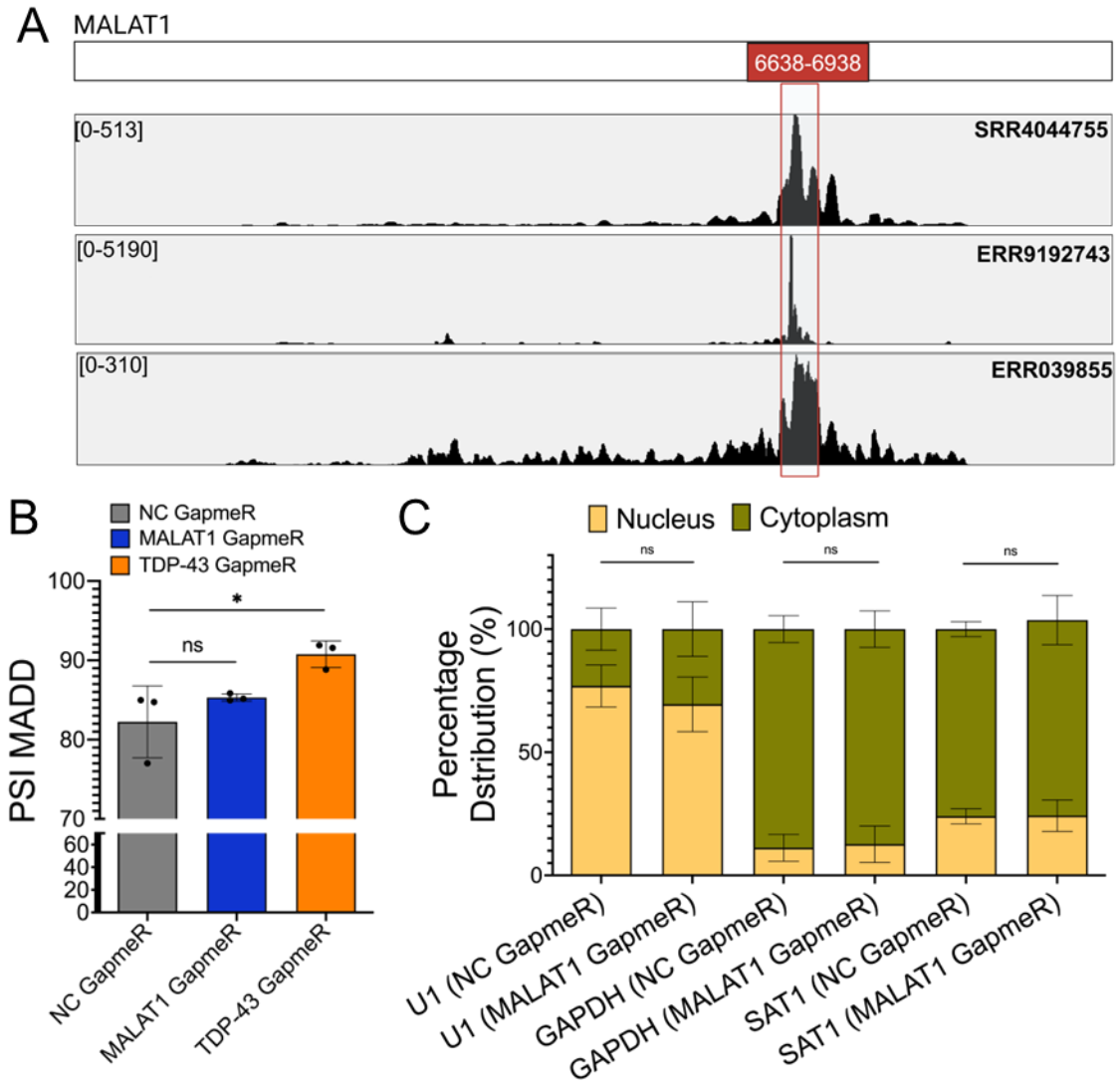
Supplemental Figure 3.1

(A) MALAT1 and TARDBP RNA expression levels upon addition of NC GapmeR, MALAT1 GapmeR, TDP-43 GapmeR, or both GapmeRs. (B) Validation of HMGB2 mRNA knockdown in HMGB2 GapmeR condition. There is no significant change in SAT1 or MADD mRNA expression levels, relative to control GAPDH mRNA. All experiments performed in supplement figure 1 are conducted in HEK293 cells. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.



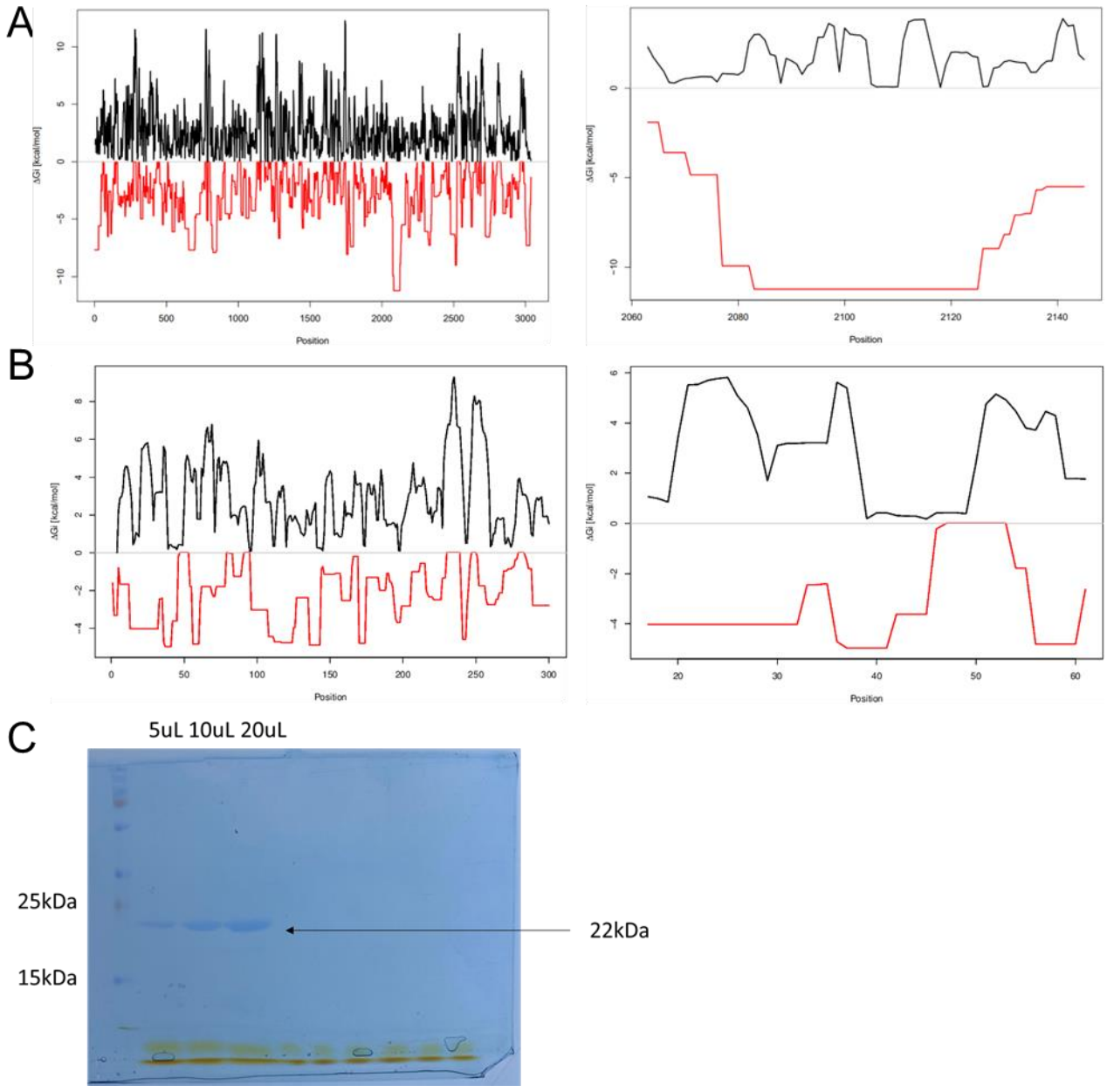
Supplemental Figure 3.2

TDP-43 protein distribution in sucrose gradient between total RNA digestion and MALAT1 GapmeR conditions. Protein localization has also been plotted as a heat map. N=3. All experiments in supplement figure 2 are conducted in K562 cells. data are plotted with error bars representing SD.



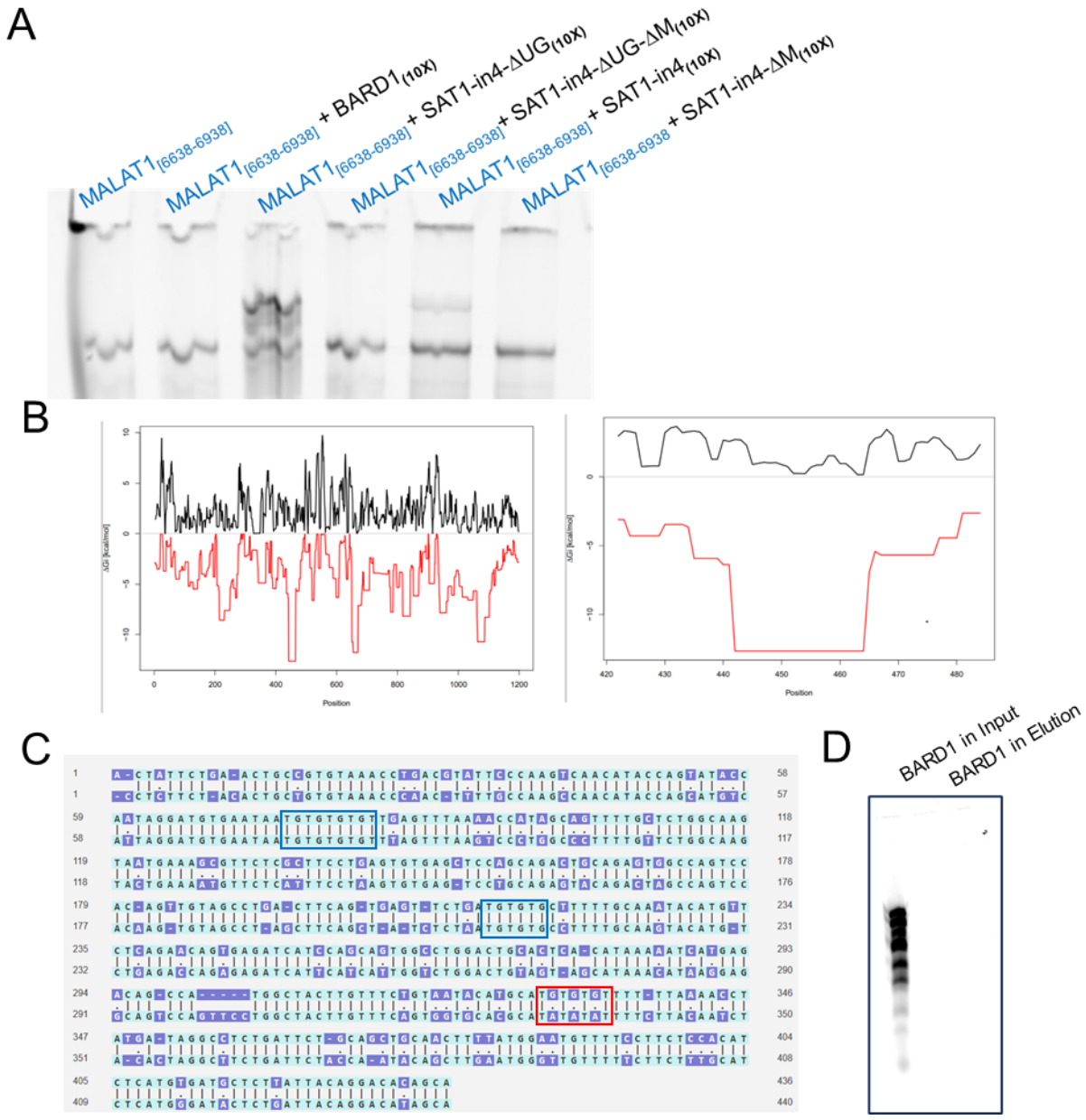
Supplemental Figure 3.3

(A) TDP-43 CLIP-seq identification of direct binding site on MALAT1 in HEK293, SH-SY5Y and H9 cells. There is a strong 300nt binding peak encompassing nucleotides 6638-6938 across all three CLIP-seq experiments. (B) Relative PSI analysis of MADD mRNA in MALAT1 and TDP-43 GapmeR conditions. C) Quantification of RNA localization in HEK293 between nucleus and cytoplasm. U1 and GAPDH RNA are used as positive controls N=3. All experiments for supplement figure 3 are conducted in HEK293 cells. All statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.



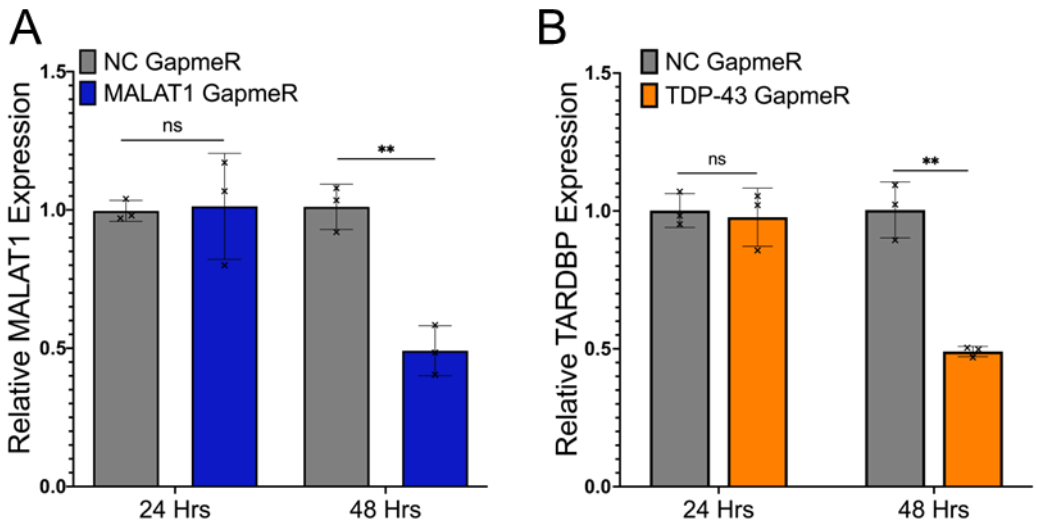
Supplemental Figure 3.4

(A) Minimal free energy hybridization prediction of MALAT1 and SAT1-in4 RNA constructs. (B) Lack of strong minimal free energy hybridization region between MALAT1 and BARD1 RNA constructs. (C) Coomassie gel showing homogenic purity of tandem TDP-43 RRM construct, and increasing linear densitometry with increasing protein concentrations.



Supplemental Figure 3.5

(A) RNA-RNA hybridization of MALAT1 with various SAT1 constructs and negative control BARD1. (B) Validation of Malat1 and mSAT1 minimal free energy hybridization regions, showing strong predictive hybridization regions. (C) Evolutionary conservation between hSAT1 and mSAT1 reveals conservation of 2 UG repeats (Black box) and non-conservation of 1 UG repeat sequence (Red box). (D) Fluorescently labelled BARD1 RNA is visible in input of TDP-43-MALAT1-BARD1 assembled complexes, but is not enriched in elution samples.



Supplemental Figure 3.6

(A) Validation of MALAT1 knockdown in MALAT1 GapmeR conditions upon 24 and 48 hours of transfection. N=3. (B) Validation of TARDBP mRNA knockdown in TDP-43 GapmeR conditions upon 24 and 48 hours of transfection. N-3. All experiments for supplement figure 6 are conducted in SH-SY5Y cells. statistical significance is conducted with student unpaired t-test with equal variance. * = $p<0.05$, ** = $p<0.01$. All data are plotted with error bars representing SD.