

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Julia Lannes de Amorim

Date

THE RNA EXOSOME: IMPLICATIONS FOR GENE EXPRESSION REGULATION AND
NEUROLOGICAL PATHOLOGY

By

Julia L. de Amorim
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Biochemistry, Cell, and Developmental Biology

Anita H. Corbett, Ph.D.
Advisor

Victor Faundez, M.D., Ph.D.
Committee Member

Yue Feng, M.D., Ph.D.
Committee Member

John R. Hepler, Ph.D.
Committee Member

Nicholas T. Seyfried, Ph.D.
Committee Member

David S. Yu, M.D., Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

The RNA Exosome: Implications for Gene Expression Regulation and Neurological Pathology

By

Julia L. de Amorim

B.A., Georgia State University, 2013

B.S., Armstrong State University, 2017

Advisor: Anita H. Corbett, Ph.D.

An abstract of a dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry, Cell, and Developmental Biology 2023

Abstract

THE RNA EXOSOME: IMPLICATIONS FOR GENE EXPRESSION REGULATION AND NEUROLOGICAL PATHOLOGY

By Julia Lannes de Amorim

The RNA exosome is a ten-subunit complex that mediates both RNA processing and degradation. This complex is evolutionarily conserved and plays many roles in regulating gene expression and protecting the genome, including modulating the accumulation of R-loops at sites of transcription. The RNA exosome interacts with specific target RNAs for decay or processing via interacting proteins termed cofactors. Although the RNA exosome complex is routinely referred to as ubiquitously expressed, little is known about the tissue- or cell-specific expression of the RNA exosome complex or any individual subunit. Recently, missense mutations in genes encoding structural subunits of the RNA exosome have been linked to a variety of distinct neurological diseases, many of them childhood neuronopathies with at least some degree of cerebellar atrophy. Understanding how these missense mutations lead to the disparate clinical presentations that have been reported for this class of diseases necessitates investigation of how these specific changes alter cell-specific RNA exosome function. One possibility for how single amino acid changes could cause neurological disease is that the RNA exosome partners with cell- or tissue-specific protein cofactors. Here, we highlight recent studies that model pathogenic variants in RNA exosome subunits. In this study, we leverage publicly available RNA-sequencing data to analyze RNA exosome subunit transcript levels in healthy human tissues, focusing on those tissues that are impacted in exosomopathy patients described in clinical reports. We additionally employed a murine neuronal cell line (N2A) and performed immunoprecipitation of the RNA exosome subunit, EXOSC3, followed by mass spectrometry to obtain a snapshot of the RNA exosome interactome. We validated an interaction with DDX1, a putative RNA helicase. DDX1 plays roles in double-strand break repair, rRNA processing, and RNA metabolism. To explore shared functions of EXOSC3 and DDX1, we investigated the interaction after inducing DNA damage, and performed DNA/RNA immunoprecipitation followed by sequencing (DRIP-Seq) and RNA-seq on N2A cells depleted of either EXOSC3 or DDX1. These findings suggest that EXOSC3 and DDX1 function together in the absence of DNA damage to modulate spontaneous events such as RNA-DNA hybrid (R-loop) formation. Taken together, these analyses suggest that specific subunits could play impactful roles that affect the RNA exosome complex in gene expression and neurological pathology.

The RNA Exosome: Implications for Gene Expression Regulation and Neurological Pathology

By

Julia L. de Amorim

B.A., Georgia State University, 2013

B.S., Armstrong State University, 2017

Advisor: Anita H. Corbett, Ph.D.

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry, Cell, and Developmental Biology 2023

Acknowledgements

I would first like to acknowledge my thesis mentor, Dr. Anita H. Corbett. Her guidance led me to become the scientist and passionate science communicator I am today. She supported my career exploration at every turn and never discouraged my enthusiasm for activities that eventually led me to my career goals. She taught me the value of communicating science and how to tailor my presentations and writings to a specific audience. She has sent me around the globe to meet other wonderful scientists and present my novel findings. I will be eternally grateful for her mentorship.

I next want to thank my thesis committee members, Drs. Victor Faundez, Yue Feng, John Hepler, Nicholas Seyfried, and David Yu. I deeply appreciate your feedback at every meeting, presentation, and exam. You are also the reason I have become the scientist and science communicator I am today. Each one of you has encouraged me to think outside the restraints of our field and urged me to continue to pursue what motivated me.

I am incredibly grateful for my lab mates, current and past. I have attained valuable mentorship from Drs. Sara Leung, Milo Fasken, Derrick Morton, Ayan Banerjee, and Kevin Morris. I have also received support from Drs. Annie McPherson, Stephane Jones, Christy Kinney, Laramie Lemon, and Celina Jones as well as current graduate students Carly Lancaster, Jordan Goldy, and Dory Fawwal. I give special thanks to my lab twin Dr. Maria Sterrett, who also entered the BCDB program among the 2017 cohort, rotated and joined the Corbett lab with me, and worked on the RNA exosome project. Our names are at least different, though many people mixed us up for the entire six-year tenure of our graduate education. I cannot conceive my time here without Maria.

I also am deeply grateful for my fiancé, David Liss. We have been together for nearly the entire duration of my time at Emory University. He has seen me push through every challenge, and I could not have persevered without his undying support.

I am thankful for my family, who luckily lives nearby and whom I was fortunate to visit often during my graduate career. They helped me unwind from deep convoluted scientific questions I attempted to answer with meals, games, vacations, and lively discussions.

Last, but certainly not least, I thank my children, Simone and Kennedy, who are cats and not human but my children all the same. They were always lending a reluctant paw, showering me with 90 second snuggles, and yowling their support for me at the top of their lungs.

CHAPTER 1: INTRODUCTION.....	1
THE RNA EXOSOME: A COMPLEX COMPLEX.	2
PATHOGENIC MISSENSE VARIANTS IN SUBUNIT GENES OF THE RNA EXOSOME ARE LINKED TO DISEASE.	3
<i>EXOSC3</i> MUTATIONS IMPAIR RNA EXOSOME FUNCTION AND ORGANISM VIABILITY.	6
SHRF-CAUSING PATHOGENIC VARIANTS IN <i>EXOSC2</i>	9
NOVEL <i>EXOSC5</i> MUTATIONS IMPAIR RNA EXOSOME ACTIVITY.....	10
CONCLUSIONS AND FUTURE PERSPECTIVES.....	11
FIGURE 1.1. PATHOGENIC MISSENSE VARIANTS IN STRUCTURAL SUBUNITS OF THE RNA EXOSOME CAUSE HUMAN DISEASE WITH DIVERSE CLINICAL PRESENTATIONS.....	16
TABLE 1.1. EXOSOMOPATHY PATHOGENIC MISSENSE VARIANT MODELS.....	17
FIGURE 1.2. MODEL OF HOW PATHOGENIC MISSENSE MUTATIONS AFFECT THE RNA EXOSOME.	18
FIGURE 1.3. A PROPOSED MODEL OF HOW PATHOGENIC MISSENSE VARIANTS IN <i>EXOSC</i> GENES COULD CONTRIBUTE TO VARIATION IN CLINICAL PRESENTATION.	20
CHAPTER 2.....	21
ABSTRACT	22
INTRODUCTION	23
THE RNA EXOSOME COMPLEX PLAYS CRITICAL ROLES IN GENE EXPRESSION IN SUBCELLULAR COMPARTMENTS.	24
CLINICAL PHENOTYPES OF EXOSOMOPATHIES INCLUDE NEUROLOGICAL DEFECTS.	25
MATERIALS AND METHODS	28
GENOTYPE-TISSUE EXPRESSION (GTEx) PROJECT	28
RESULTS AND DISCUSSION.....	29
A COMPARATIVE ANALYSIS OF RNA EXOSOME SUBUNITS REVEALS DISPARATE REQUIREMENTS OF TRANSCRIPTS IN TISSUES.....	29
FIGURE 2.1. THE RNA EXOSOME SUBUNIT TRANSCRIPTS ARE UBIQUITOUSLY EXPRESSED IN HUMAN TISSUE.	34
FIGURE 2.2. RNA EXOSOME SUBUNIT TRANSCRIPT LEVELS ARE HIGH IN THE CEREBELLAR HEMISPHERE/CEREBELLUM COMPARED WITH OTHER TISSUES.	36
CHAPTER 3.....	37
ABSTRACT	38
INTRODUCTION	39
RESULTS	42
PROTEOMICS REVEAL A SUITE OF <i>EXOSC3</i> INTERACTORS.....	42
PUTATIVE HELICASE <i>DDX1</i> INTERACTS WITH THE RNA EXOSOME IN THE NUCLEUS.	43

THE INTERACTION BETWEEN EXOSC3 AND DDX1 DECREASES IN RESPONSE TO DNA DAMAGE.....	45
DEPLETION OF EXOSC3 OR DDX1 RESULTS IN RRNA PROCESSING DEFECTS.	46
R-LOOPS ARE GLOBALLY REDUCED UPON DEPLETION OF EXOSC3 OR DDX1.....	48
DISCUSSION	51
CHAPTER 3 FIGURES	59
FIGURE 3.1. RNA EXOSOME SUBUNITS CO-IMMUNOPRECIPITATE WITH TAGGED EXOSC3.	59
FIGURE 3.2: NOVEL EXOSC3/RNA EXOSOME INTERACTORS IDENTIFIED USING LIQUID CHROMATOGRAPHY COUPLED WITH TANDEM MASS SPECTROMETRY (LC-MS/MS).	61
FIGURE 3.3: DDX1 CO-IMMUNOPRECIPITATES WITH EXOSC3.....	64
FIGURE 3.4: THE INTERACTION BETWEEN EXOSC3 AND DDX1 IS SENSITIVE TO DNA DAMAGE.....	65
FIGURE 3.5: EXOSC3 AND DDX1 ARE ROBUSTLY DEPLETED BY siRNA-MEDIATED KNOCKDOWN IN N2A CELLS.	67
FIGURE 3.6: DEPLETION OF EXOSC3 OR DDX1 RESULTS IN MISPROCESSING OF RRNA PRECURSORS.	70
FIGURE 3.7: DRIP-SEQ REVEALS THAT DEPLETION OF EXOSC3 OR DDX1 ALTERS R-LOOP REGIONS.....	71
FIGURE 3.8: RNA-SEQ SHOWS THAT DEPLETION OF EXOSC3 OR DDX1 RESULTS IN MORE SHARED DECREASED TRANSCRIPTS THAN SHARED INCREASED TRANSCRIPTS.	73
FIGURE 3.9. FILTERING DRIP READS THROUGH RNA SEQUENCING REVEALED GENES THAT ARE SIMULTANEOUSLY AFFECTED BY DEPLETIONS OF EXOSC3 OR DDX1.	75
TABLE 3.S1.....	76
FIGURE 3.S1. EXOSC3 CUSTOM-MADE ANTIBODY IS SPECIFIC, CELLULAR FRACTIONATION IS SUFFICIENT, AND THE INTERACTION BETWEEN EXOSC3 AND DDX1 IS IMPACTED BY LOSS OF RNA OR DNA.	77
FIGURE 3.S2. THE QUALITY OF THE PURIFIED RNA USED FOR THE NORTHERN BLOTS, RNA-SEQUENCING, AND DRIP-SEQUENCING ASSESSED BY 1% AGAROSE GEL AND HIGH SENSITIVITY SCREENTAPE ASSAY.....	79
FIGURE 3.S3. IGV FOR R-LOOP REGIONS IN BAMBI IS ELEVATED IN CELLS DEPLETED OF EXOSC3 OR DDX1.....	81
CONCLUSIONS AND FUTURE DIRECTIONS.....	83
EXOSOMOPATHIES AND THE BIOLOGICAL CONSEQUENCES OF PATHOGENIC MISSENSE MUTATIONS IN RNA EXOSOME SUBUNIT GENES.	85
NOVEL INTERACTIONS WITH THE RNA EXOSOME IN NEURONAL CELLS.	87
FIGURE 4.1. MODEL OF HOW THE RNA EXOSOME MAY IMPACT GENE EXPRESSION REGULATION AND NEUROLOGICAL PATHOLOGY.....	94
REFERENCES.....	96

Chapter 1: Introduction

Modeling Pathogenic Variants in the RNA Exosome

The following chapter has been published:

de Amorim J, Slavotinek A, Fasken MB, Corbett AH, Morton DJ. Modeling Pathogenic Variants in the RNA Exosome. *RNA Dis.* 2020;7:e1166. PMID: 34676290; PMCID: PMC8528344.

Chapter 1 has been modified in this dissertation from the original publication to reflect novel findings in the field.

The RNA exosome: a complex complex.

The genome is identical in all somatic cells and comprises thousands of genes, many of which are expressed as RNA transcripts. This extensive primary transcriptome is converted into functional RNAs, and each cell has specific transcriptomic requirements. To achieve this, RNAs must be precisely processed and/or degraded at the appropriate time of development and in the correct cell type to maintain cellular function. The RNA processing and degradation program in cells involve numerous RNases and processing factors. A key cellular RNA processing/decay machine is the RNA exosome. The RNA exosome is a 10-subunit complex responsible for essential RNA processing/degradation in both the cytoplasm and the nucleus. The ribonuclease activity of the RNA exosome is critical for both RNA quality control and precise processing of key RNAs, including ribosomal RNA (rRNA) (1). As shown in **Figure 1.1**, the 10 subunits of this complex are organized into a non-catalytic cap composed of three subunits (EXOSC1-3), a barrel-shaped non-catalytic core composed of six subunits (EXOSC4-9), and one catalytic 3'-5' exo/endoribonuclease subunit that sits at the base of the core (DIS3) (2-6). Most target RNAs are threaded through the cap and the central channel of the barrel to reach DIS3 for processing and/or degradation (7,8). The RNA exosome is evolutionarily conserved, and all subunits analyzed are essential in any model organisms where studies have been performed (9-16).

Genes encoding subunits of the RNA exosome complex were initially discovered in a genetic screen for rRNA processing mutants in budding yeast (1,17). Studies in *S. cerevisiae* demonstrate that 1) each subunit is essential for survival, 2) the RNA exosome processes and degrades target transcripts in a 3'-5' orientation, and 3) conditional mutations in genes encoding the RNA exosome subunits impair RNA metabolism (1,13,17,18). Early structures of the RNA exosome from a number of organisms provided key insight into how this complex could both

process and decay RNA (19-21). Since then, a number of elegant structural and *in vitro* biochemical studies have been employed to understand the many functions of this critical complex (22,23).

The RNA exosome regulates/processes several classes of RNAs in different cellular compartments (24-27). In the nucleolus, RNA exosome-mediated processing is essential for the production of mature rRNA (24). Within the nucleus, this complex also processes and/or degrades small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), tRNAs, cryptic unstable transcripts (CUTs) in yeast and promoter-upstream transcripts (PROMPTs) in mammals (18,25,28-30). In the cytoplasm, key targets include normal mRNAs in the turnover pathway and aberrant mRNA transcripts, such as those lacking a stop codon, in quality control pathways (31). The RNA exosome also continues to trim rRNA for precise maturation in the cytoplasm (32). In addition, the RNA exosome regulates the levels of a variety of different transcripts (2,24,25). To recognize and process/degrade distinct targets, the RNA exosome interacts with cofactors, proteins that associate with the complex (33). Several RNA exosome cofactors that serve as RNA helicases, scaffolds, additional ribonucleases, and polyadenylases have been described (2,5-7,28,34-37). Cofactors have been primarily characterized in budding yeast, but more recent studies have identified mammalian cofactors (2,27,38,39), providing fundamental insights into RNA exosome specificity for processing and decay of target RNAs.

Pathogenic missense variants in subunit genes of the RNA exosome are linked to disease.

Although the function of the RNA exosome is essential (17,40), a number of studies have now identified mutations in genes encoding structural subunits of this complex linked to diverse clinical presentations (**Table 1.1**). These diseases are termed *exosomopathies* (10). The initial

report linking RNA exosome genes to disease described several pathogenic variants in *EXOSC3* that cause pontocerebellar hypoplasia type 1B (PCH1B) (16). Subsequent studies have linked *EXOSC1* (41), *EXOSC2* (42), *EXOSC5* (43), *EXOSC8* (9), and *EXOSC9* (10) to a variety of clinical presentations (44).

Although clinical presentations of exosomopathies are variable, effects on the cerebellum are a common feature. In several of the exosomopathies described to date, patients present with developmental delay of the cerebellum (cerebellar hypoplasia or pontocerebellar hypoplasia (PCH)) with progressive degeneration of the cerebellum. The cerebellar pathology is diverse and is typically associated with additional clinical manifestations (44). Mutations in *EXOSC1* are linked to PCH1F (41). Mutations in *EXOSC3* give rise to PCH1B, a disease characterized by atrophy of the cerebellum and the pons (16,45). *EXOSC8* mutations cause PCH1C, characterized by hypomyelination with spinal muscular atrophy and cerebellar hypoplasia (9). *EXOSC9* mutations give rise to PCH1D, a spinal motor neuronopathy coupled with cerebellar atrophy (10,46). While only a few patients with mutations in *EXOSC5* have been described, these patients also show cerebellar abnormality as a common clinical feature (47). In contrast to the other *EXOSC* mutations, pathogenic variants in *EXOSC2* only cause mild/borderline cerebellar atrophy and patients present with short stature, hearing loss, retinitis pigmentosa, and distinctive facies (denoted as SHRF) (42). Whether the cerebellar hypoplasia and atrophy observed in the exosomopathies are part of a clinical spectrum that results from the same pathological process and molecular mechanism or distinct manifestations is still unknown.

Many pathogenic mutations expressed in patients thus far result in single amino acid substitutions. These substitutions often occur in conserved protein domains within the subunit (**Figure 1.1B**). The three cap subunits, *EXOSC1*, *EXOSC2*, and *EXOSC3*, contain conserved N-

termini and S1 RNA-binding domains. EXOSC2 and EXOSC3 contain an additional conserved KH RNA-binding domain. Within the three cap subunits, single amino acid changes in patients have been reported to occur in the N-terminal domains (44). The core subunits share a PH domain. The three subunits that have thus far been linked to disease, EXOSC5, EXOSC8, and EXOSC9, have all been reported to incur at least one single amino acid change (44). Despite these single amino acid changes occurring in conserved protein domains, the clinical outcomes observed in patients are diverse.

Little is understood about why these pathogenic missense variants in genes encoding structural subunits of an essential complex that is ubiquitously expressed give rise to a broad range of clinical presentations. All patients with exosomopathies described thus far have at least one missense variant in an *EXOSC* gene (**Table 1.1**) (44). Some patients are homozygous for the same missense variant, others are compound heterozygous for different missense variants, and some patients have a missense variant inherited *in trans* to a deletion or loss-of-function variant. The complete loss of the RNA exosome is lethal (9-16); therefore, the missense variants likely provide residual RNA exosome function in all patients. These specific variants in different *EXOSC* genes may underlie the disparate clinical presentations of patients. The pathogenic amino acid changes (**Figure 1.2A**) could the integrity of the RNA exosome complex (**Figure 1.2B**), alter the functional interactions with proteins (termed cofactors) (**Figure 1.2C**), ultimately affecting downstream RNA targets (**Figure 1.2D**).

Studies to define the molecular mechanisms underlying pathology in exosomopathies have used several approaches: 1) immortalized patient cells; 2) deletion or depletion of the affected *EXOSC* subunit; and 3) modeling of the pathogenic missense variants in either model genetic systems or cultured cells. Ultimately, understanding how defects in RNA exosome function

contribute to disease pathology will require studies investigating how the pathogenic amino acid substitutions impact the function of the complex. At this time, *in vivo* studies that model missense mutations to understand how disease-linked amino acid changes could alter RNA exosome function have employed budding yeast and *Drosophila melanogaster*. Here, we highlight recent studies that model pathogenic missense variants in *EXOSC3* (11,12,14), *EXOSC2* (48) and *EXOSC5* (47).

***EXOSC3* mutations impair RNA exosome function and organism viability.**

Initial studies to explore the functional consequences of pathogenic variants in *EXOSC3* employed the budding yeast model system (11,12). The first observation from these studies is that pathogenic variants in *EXOSC3* modeled in yeast did not severely impact yeast cell growth or viability. This result is perhaps not surprising as each subunit of the RNA exosome is essential in the systems where this has been tested (1,9-16,49). The presumption is that changes that significantly impair the function of this essential complex, which might confer a significant growth defect in yeast cells, may not be compatible with human development. Although severe growth defects were not observed, both these studies showed that the yeast *EXOSC3* variant corresponding to *EXOSC3* W238R (W195R in budding yeast *EXOSC3*) conferred a temperature-sensitive growth defect when expressed as the sole copy of yeast *EXOSC3* (11,12). Moreover, these cells showed a significant impact on RNA processing and degradation mediated by the RNA exosome. Yeast *EXOSC3* variants corresponding to *EXOSC3* G31A (G8A in yeast *EXOSC3*), G191C (G148C in yeast *EXOSC3*), and W238R (W195R in yeast *EXOSC3*), showed impaired rRNA processing with the most profound effects evident for the yeast W195R variant (11). Additionally, misprocessing and accumulation of several RNA exosome targets, including CUTs and pre-

snRNA, occur in cells expressing the W195R variant as the sole copy of yeast EXOSC3 (12). In contrast, no effect on cytoplasmic RNA exosome function was detected (12). The yeast model studies thus indicate that the W195R variant impairs cell function and significantly affects RNA processing, suggesting that the EXOSC3 W238R variant is deleterious. This allele has only been identified in the compound heterozygous state in patients (16,45), raising the possibility that EXOSC3 W238R may not confer sufficient RNA exosome activity to support life as a homozygous variant.

To begin to address how pathogenic amino acid substitutions could impair RNA exosome function, one study compared both the steady-state levels and stability of the yeast EXOSC3 variants to wild-type EXOSC3 (12). Results of this analysis demonstrated that the yeast W195R variant protein is relatively stable when expressed as the sole copy of the yeast EXOSC3 protein but becomes destabilized under conditions where a wild-type copy of the subunit is also present. These results suggest that perhaps the pathogenic subunits are not as efficiently incorporated into the complex as the wild-type subunits, or once pathogenic subunits are incorporated, the complex is not as stable. These yeast studies were complemented by analyzing mouse EXOSC3 variant protein levels in cultured mouse neuronal cells. This analysis showed that the steady-state level of the mouse EXOSC3 tagged protein modeling the W238R variant is reduced in these cells compared to the wild-type tagged EXOSC3 (12). Thus, a decrease in overall complex level could contribute to pathology in exosomopathies as suggested from analyses of other EXOSC variants, such as EXOSC9 (10); however, reconciling the very diverse clinical presentations of these diseases with a simple loss of or decrease in overall complex function is difficult.

Beyond affecting protein levels, pathogenic amino acid substitutions could also alter key interactions with other RNA exosome subunits or with the associated cofactors. Indeed, one study

demonstrated that the yeast EXOSC3 W195R variant corresponding to EXOSC3 W238R shows decreased affinity for the scaffolding cofactor, MPP6, compared to wild-type yeast EXOSC3 (6). These results support a model where altered interactions with RNA exosome cofactors could contribute to disease pathology.

Studies modeling pathogenic variants in *EXOSC3* have been extended to *Drosophila*. This system enables functional consequences of EXOSC3 variants within the nervous system and brain to be studied in a genetically tractable system. A previous study in *Drosophila* demonstrated that RNA exosome subunits are essential in flies (50), consistent with the results obtained in budding yeast (1). In a more recent study, CRISPR/Cas9 genome editing was used to engineer pathogenic variants of *EXOSC3* into the *Drosophila* genome (14). This is the first study that analyzes RNA exosome mutations recapitulated at the genome level in a multi-tissue organism. The study modeled three patient genotypes (16,45): homozygous G31A; homozygous D132A; and D132A over a deficiency to model patients heterozygous for the D132A pathogenic variant inherited *in trans* to a deletion in *EXOSC3*. Results of this analysis show a striking genotype-phenotype correlation with respect to fly viability, lifespan, and locomotor function. The pathogenic variants that are most severe in patients (51) correlate with those that cause the most striking phenotypes in flies. These mutant *EXOSC3* flies show morphological defects in the mushroom body, the area of the fly brain that controls learning and memory (52), which also correlate with the severity of the different *EXOSC3* alleles modeled. Finally, RNA sequencing of the heads of these mutant flies revealed an increase in the steady-state levels of a number of important neuronal transcripts, a result that is consistent with the role of this complex in RNA decay. This study developed a multi-cellular model to explore the consequences of pathogenic RNA exosome variants and provided insight into target RNAs affected in this model (14).

SHRF-causing pathogenic variants in EXOSC2.

Mutations in *EXOSC2* give rise to a novel syndrome characterized by short stature, hair loss, retinitis pigmentosa, and distinctive facies (SHRF) (42). A recent study combined analysis of patient samples, biochemical approaches, and studies in *Drosophila* to explore the functional consequences of pathogenic missense variants in *EXOSC2* (48). Biochemical analyses of patient lymphoblasts, transfected HEK293T cells, and cultured keratinocytes demonstrated that the G198D variant, but not the G30V variant, affects *EXOSC2* protein stability and interactions with other RNA exosome components. While these authors did not create a fly model of the pathogenic variants in *EXOSC2*, they did test whether the *EXOSC2* G30V variant could rescue defects observed in the eye in rare “escapers” where fly *EXOSC2* was deleted. The eye defect was partially rescued when the human wild-type *EXOSC2* gene was expressed, but not the pathogenic variant, providing evidence that *EXOSC2* G30V does not retain the function of wild-type *EXOSC2* (48). RNA-sequencing of patient samples identified several dysregulated autophagy pathway genes (48). In addition, in patient-derived B-lymphoblast cells with mutations in *EXOSC2*, overall RNA exosome subunit abundance is reduced and the *EXOSC2* protein is unstable (48). These findings are consistent with other studies that show a decrease in RNA exosome subunit levels in patient-derived samples (10), but do not readily explain why patients with mutations in different *EXOSC* genes display such a variety of clinical presentations as they used cell types in this study that are not pathologically represented in exosomopathy patients.

Novel *EXOSC5* mutations impair RNA exosome activity.

A recent study reported five patients with biallelic variants in the *EXOSC5* gene (47). Three of the four patients who learned to walk showed ataxia, and four of the five patients' brain imaging showed hypoplasia of the cerebellum or cerebellar vermis (47). This study employed three approaches to examine the link between *EXOSC5* and disease pathology; two of these approaches modeled the pathogenic variants that have been identified in *EXOSC5*. The initial approach employed zebrafish to assess the requirement for *EXOSC5* in neurodevelopment has also been employed for the analysis of *EXOSC3* (16), *EXOSC8* (9), and *EXOSC9* (10). The CRISPR-Cas9 system was employed to generate an allele with predicted loss of *EXOSC5* function. Consistent with previous studies of other RNA exosome subunit genes, zebrafish lacking *EXOSC5* showed profound defects in growth, developmental, and brain morphology (47). To extend this analysis and explore the functional consequences of pathogenic variants in *EXOSC5*, the missense variants identified in *EXOSC5* (I114T, M148T, and L206H) were all modeled in budding yeast. Only the yeast *EXOSC5* variant corresponding to *EXOSC5* L206H (L191H in yeast *EXOSC5*) showed growth defects, manifested as temperature-sensitive growth. Immunoblotting showed no statistically significant change in yeast *EXOSC5* protein levels for any of these variants. Consistent with the growth defect observed, the yeast variant corresponding to *EXOSC5* L206H caused defects in U4 snRNA and 7S pre-rRNA processing (47). Biochemical studies in cultured mouse neuronal cells were performed to explore the interactions of mouse *EXOSC5* variants with other subunits of the RNA exosome. Interestingly, the *EXOSC5* L206H and I114T variants showed decreased interaction with multiple other subunits of the complex. No defect in interaction was detected for the *EXOSC5* M148T raising the question of how this amino acid substitution contributes to pathology and suggests different variants have distinct pathological mechanisms.

Conclusions and Future Perspectives

The model organisms used to study RNA exosome biology have shed light on the complicated cellular roles of this multi-functional protein complex. Thus far, pathogenic missense variants in *EXOSC* genes that encode structural RNA exosome subunits have been linked to diverse clinical presentations with the majority causing some degree of cerebellar hypoplasia and/or atrophy. These *EXOSC* variants appear to have tissue-specific consequences. An ideal system to further explore the consequences of these pathogenic variants would be in an affected patient tissue or cultured primary cells. These samples are difficult to obtain because the disease is rare with only small numbers of patients identified to date. The cerebellum consists largely of Purkinje and granule cells (53); therefore, an ideal *in vitro* system would be a genome-edited cerebellar cell type organoid because the three-dimensional shape allows for multiple cell types. *In vivo* studies in model organisms thus far have allowed for simple recapitulation of pathogenic variants. Budding yeast are advantageous because they are simple to use, and multiple molecular functions of the conserved RNA exosome complex can be readily assayed. However, yeast lack relevant cell or tissue types. *Drosophilae* are multi-tissue organisms with a complex nervous system and a brain, which have been used extensively to model human disease (54). However, fly brain structures lack elements of the human brain, such as a cerebellum, and some neurons are not well conserved (55).

A mammalian system would provide further insight into how specific pathogenic amino acid variants alter the function of the RNA exosome and cause diverse biological changes that underlie pathology. Thus far, no whole organism mouse models for mutations in any RNA exosome gene have been described. One study employed an *ex vivo* approach to swap exons 2 and

3 within *EXOSC3* in B cells (56). The authors exploited this *ex vivo* system to explore the downstream effects of the loss of this RNA exosome subunit in B cells and identify key RNA exosome targets in this cell type (56). A transgenic mouse that uses genome editing to incorporate pathogenic missense variants of *EXOSC* genes into the genome would greatly advance research in this field. Currently, the Jackson Laboratory site (57) lists CRISPR-generated knockout mouse strains for *EXOSC1* and *EXOSC2*, but there are no reports of researchers attempting to derive or analyze these strains.

A proposed model suggests that depletion of a structural subunit of the RNA exosome in model organisms recapitulates the patient disease because immunoblots of patient fibroblasts, myoblasts, and skeletal muscle samples show an overall steady-state decrease in protein levels of the affected subunits and other subunits in the complex (9,10). To date, no biochemical studies using cells from cerebellum have contributed to this model, likely due to limited sample availability. However, this model is difficult to reconcile with patients that show diverse clinical presentation. Some of the *EXOSC* pedigrees identified to date also call into question whether the primary driver of pathology is a decrease in RNA exosome subunit or complex levels. Several of the pedigrees for patients include seemingly unaffected parents with presumed loss of function of one *EXOSC* allele (16,44,47,51). Consistent with this idea, flies heterozygous for a deficiency that removes the *Drosophila EXOSC3* gene do not show any of the phenotypes noted in the disease model flies (14). These flies appear similar to wild-type flies in all assays performed, showing that even with presumably only 50% of expression of the *EXOSC3* gene, no phenotype is detected. Likely, pathology results from some combination of a decrease in overall RNA exosome levels and consequences of the specific pathogenic variants present. Further studies that examine the

pathogenic missense variants of the *EXOSC* genes as well as the full RNA exosome complex are required to fully understand the molecular defects that contribute to pathology.

Experimental systems that deplete or delete an *EXOSC* gene provide critical insight but also simplify a more complex story. A major challenge remains understanding why the clinical presentation of patients with mutations in genes that encode a single complex show such variable pathology with some overlapping, but some distinct tissues affected. At the mechanistic level, pathogenic amino acid changes could, and certainly do (10), decrease the overall levels of the individual RNA exosome subunit and/or the complex. The changes might also alter the function of the complex at a molecular level, disrupting RNA binding or interactions with other subunits or cofactors. Moreover, certain cell types may express specific RNA exosome cofactors that interact and stabilize the RNA exosome, whereas other affected tissues may not express these cofactors (**Figure 1.3**). Loss of interactions with these cofactors could contribute to altered stability of the complex or changes in target specificity. These potential mechanistic consequences would all affect exosome-mediated RNA processing/decay.

New pathogenic *EXOSC* alleles are being reported by physicians worldwide. To date, pathogenic variants in *EXOSC1*, *EXOSC4*, *EXOSC6*, and *EXOSC7* have not been reported (**Table 1.1**); however, this is likely to change. GeneMatcher®, a freely available website designed to enable connections between clinicians and researchers (58), will likely continue to lead to identification of new alleles in these genes and in new genes as recently illustrated in the collaborative study of *EXOSC5* (47). As additional pathogenic variants are identified, this could provide insight into whether there is a common mechanism underlying pathology. A combination of approaches that includes functional studies in model organisms and analysis of patient samples will be critical to understand the mechanisms that underlie pathology. Further insight into the

molecular consequences of single amino acid changes in *EXOSC* genes via genome-edited model systems will be required to paint the full canvas of RNA exosome biology and disease.

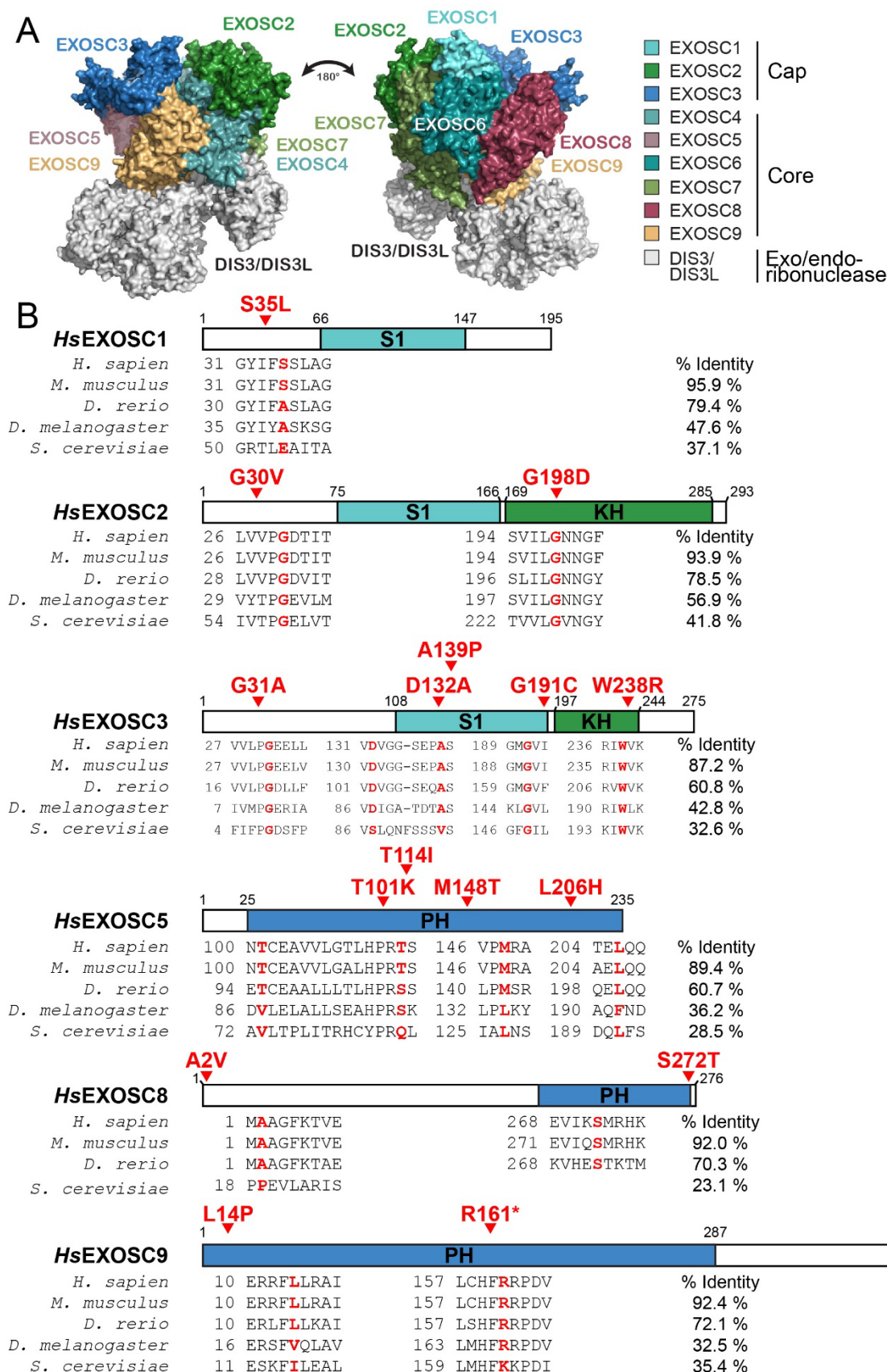

























Figure 1.1. Pathogenic missense variants in structural subunits of the RNA exosome cause human disease with diverse clinical presentations.

(A) A structural model of the 10-subunit human RNA exosome is shown (PDB #6H25) (4). The cap comprises three subunits (EXOSC1, EXOSC2, and EXOSC3), the hexameric core comprises six subunits (EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, and EXOSC9), and the catalytic base is an exo/endoribonuclease (DIS3). (B) The domain structures of the subunits altered in disease are shown: EXOSC1, EXOSC2, EXOSC3, EXOSC5, EXOSC8, and EXOSC9. Sequence alignments of human (*H. sapiens*), mouse (*M. musculus*), zebrafish (*D. rerio*), fruit fly (*D. melanogaster*), and yeast (*S. cerevisiae*) orthologs are depicted below the structures to highlight the conserved residues altered in disease and the flanking conserved regions. Numbers preceding the sequences indicate the amino acid position and the red arrows above the domain structures indicate the location of the respective amino acid substitution. The overall percent identity of the EXOSC orthologs compared to human EXOSC proteins is shown to the right of the sequence alignments. Core subunit EXOSC8 *D. melanogaster* ortholog does not exist and therefore is not shown. Furthermore, the orthologous sequence of the missense mutation resulting in S272T in EXOSC8 for yeast does not exist and thus the alignment is not pictured.

Table 1.1. Exosomopathy Pathogenic Missense Variant Models

Structural subunit gene	Genotype ⁱ	Pathogenic variants	Exosomopathy	Pathogenic Variant Models
<i>EXOSC1</i>	Homozygous	S35L	PCH1F (41)	
<i>EXOSC2</i>	Homozygous	G30V	SHRF (42) ⁱⁱ	 
	Heterozygous	G30V/G198D	SHRF	 ⁱⁱⁱ 
<i>EXOSC3</i>	Homozygous	D132A	PCH1B (11,16)	 
	Homozygous	G31A	PCH1B	  
	Homozygous	G191C	PCH1B	
	Heterozygous	G31A/W238R	PCH1B	 ^{iv}  ^v
	Heterozygous	D132A/del ^{vi}	PCH1B	
<i>EXOSC4</i>				
<i>EXOSC5</i>	Homozygous	L206H	Novel exosomopathy	 
	Homozygous	M148T	Novel exosomopathy	 
	Heterozygous	T114I/del ^{vi}	Novel exosomopathy	 
<i>EXOSC6</i>				
<i>EXOSC7</i>				
<i>EXOSC8</i>	Homozygous	S272T	PCH1C (9)	
	Homozygous	A2V	PCH1C	
<i>EXOSC9</i>	Homozygous	L14P	PCH1D (51)	

*D. melanogaster**S. cerevisiae*

Cultured cell lines



Patient samples/cultured patient cells

ⁱ Homozygous or compound heterozygousⁱⁱ Short stature, hair loss, retinitis pigmentosa, distinct facies (SHRF)ⁱⁱⁱ HEK293T cells were transfected with either EXOSC3 G30V or the EXOSC2 G198D variant^{iv} Haploid budding yeast either expressed the yeast variant corresponding to EXOSC3 G31A or EXOSC3 W238R^v Neuro2A cells were transfected with the variant corresponding to either EXOSC3 G31A or EXOSC3 W238R^{vi} Genetic deletion

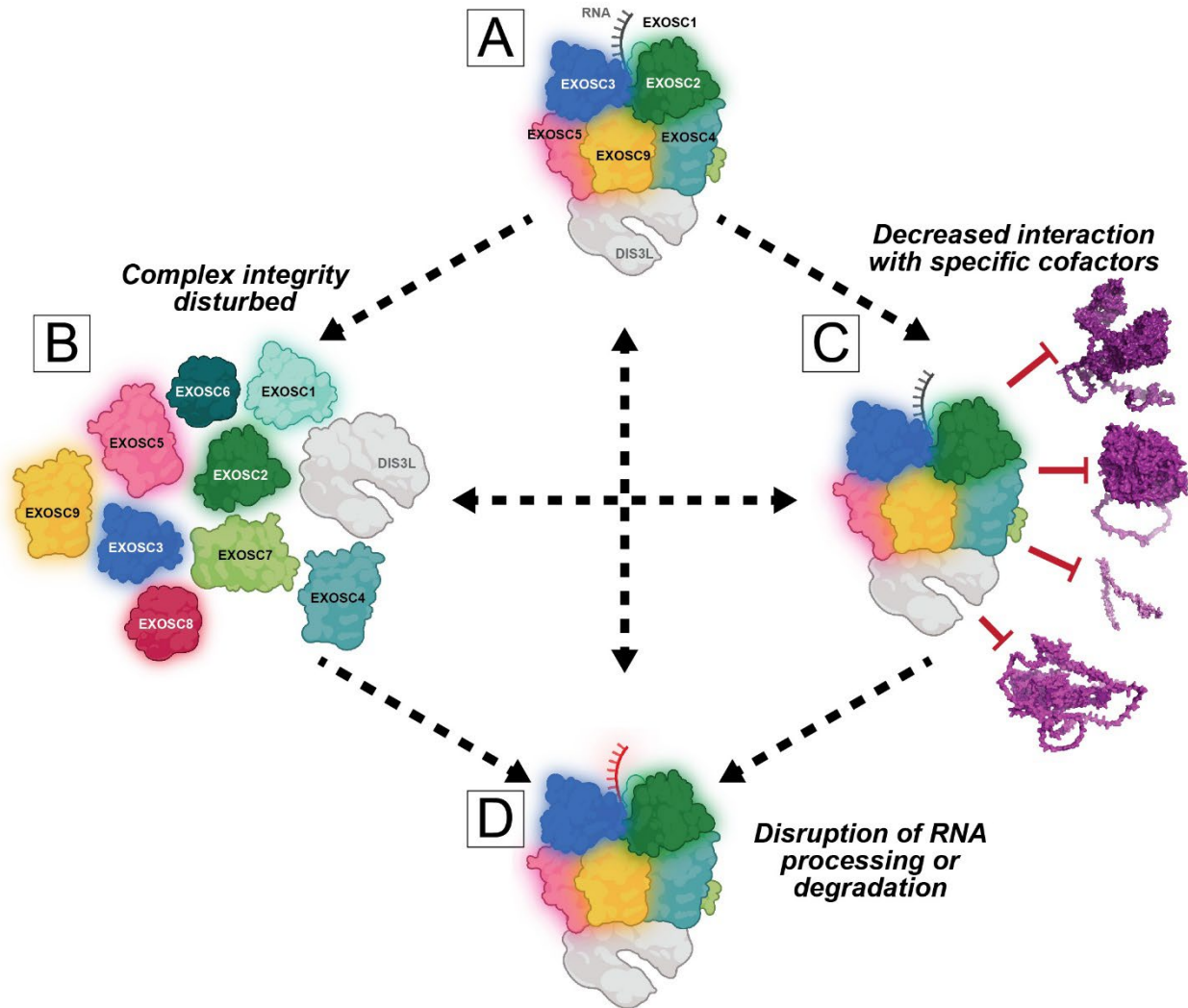


Figure 1.2. Model of how pathogenic missense mutations affect the RNA exosome.

(A) To date, missense mutations in RNA exosome subunit genes linked to human disease are reported in *EXOSC1*, *EXOSC2*, *EXOSC3*, *EXOSC5*, *EXOSC8*, and *EXOSC9* (glowing). A number of consequences could arise due to single amino acid changes in the subunits. (B) A single amino acid change could impact the integrity of the assembly of the RNA exosome complex, (C) may disrupt the proper processing and/or degradation of specific RNA targets, or (D) decrease the interactions with cofactors such as helicases, scaffolds, and other ribonucleases. Any of the consequences described could affect the other. The impeded complex assembly may disrupt RNA processing and cofactor interactions. The decreased interactions with cofactors may impede

complex assembly and disrupt RNA processing. The disruption of RNA processing/degradation could obstruct the gene expression of cofactors or assembly factors.

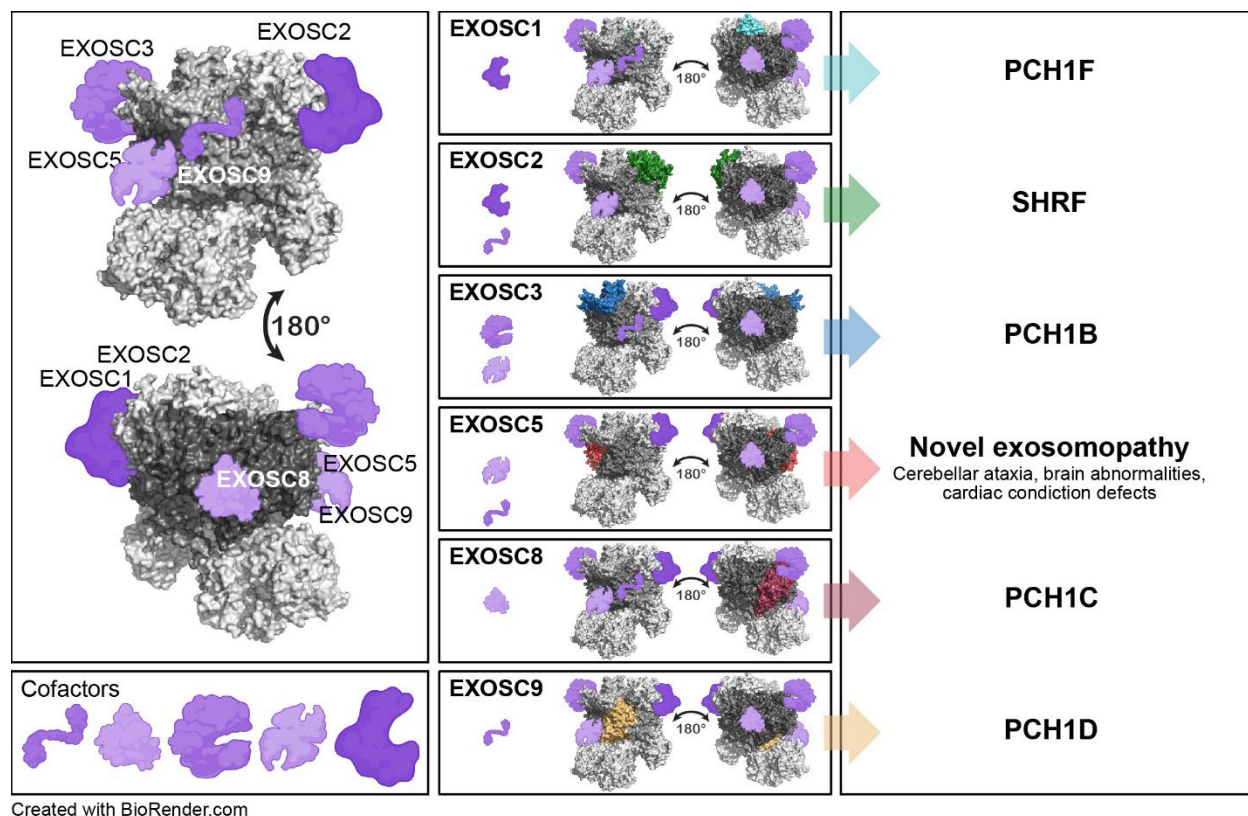


Figure 1.3. A proposed model of how pathogenic missense variants in *EXOSC* genes could contribute to variation in clinical presentation.

The RNA exosome interacts with specific cofactors to confer specificity for target transcripts (left). Most cofactors have been identified and studied in budding yeast or cultured cells. Different pathogenic variants in structural subunits of the RNA exosome could disrupt interactions with a subset of cofactors. These cofactors could be critical in the cerebellum or other tissues susceptible to pathology. The position of the specific subunit in the complex may contribute to specific cofactor interactions. When an amino acid substitution in a subunit arises from a missense mutation, these interactions may be perturbed. The loss of an interaction with a specific cofactor could result in distinct phenotypic consequences and therefore explain the pathological differences seen in exosomopathy patients.

Chapter 2

Analysis of RNA Exosome Subunit Transcript Abundance Across Tissues Implications for Neurological Disease Pathogenesis

The following chapter is under review at *RNA & Disease*. The preprint is available:

de Amorim, J. L., Asafu-Adjaye, D. G., & Corbett, A. H. (2023). Analysis of RNA Exosome Subunit Transcript Abundance Across Tissues Implications for Neurological Disease Pathogenesis. *bioRxiv*, 2023-06.

This work was completed in collaboration with Don G. Asafu-Adjaye, an undergraduate mentee of Julia L. de Amorim.

Abstract

Exosomopathies are a collection of rare diseases caused by mutations in genes that encode structural subunits of a ribonuclease complex termed the RNA exosome. The RNA exosome mediates both RNA processing and degradation of multiple classes of RNA. This complex is evolutionarily conserved and required for fundamental cellular functions, including rRNA processing. Recently, missense mutations in genes encoding structural subunits of the RNA exosome complex have been linked to a variety of distinct neurological diseases, many of them childhood neuronopathies with at least some cerebellar atrophy. Understanding how these missense mutations lead to the disparate clinical presentations that have been reported for this class of diseases necessitates investigation of how these specific changes alter cell-specific RNA exosome function. Although the RNA exosome complex is routinely referred to as ubiquitously expressed, little is known about the tissue- or cell-specific expression of the RNA exosome complex or any individual subunit. Here, we leverage publicly available RNA-sequencing data to analyze RNA exosome subunit transcript levels in healthy human tissues, focusing on those tissues that are impacted in exosomopathy patients described in clinical reports. This analysis provides evidence to support the characterization of the RNA exosome as ubiquitously expressed with transcript levels for the individual subunits that vary in different tissues. However, the cerebellar hemisphere and cerebellum have high levels of nearly all RNA exosome subunit transcripts. These findings could suggest that the cerebellum has a high requirement for RNA exosome function and potentially explain why cerebellar pathology is common in RNA exosomopathies.

Introduction

Exosomopathies are a collection of rare congenital pediatric diseases resulting from missense mutations in genes encoding structural subunits of the RNA exosome complex. The RNA exosome is a ribonuclease complex required for multiple critical cellular functions that dictate gene expression and post-transcriptional regulation. One of the most well-defined fundamental roles the RNA exosome plays in gene expression is mediating the precise processing of ribosomal RNA (rRNA) required to produce mature ribosomes (1,17). In addition to the maturation of rRNA, the RNA exosome targets many classes of RNAs for processing, degradation, and turnover (33).

The RNA exosome complex comprises ten subunits: nine structural noncatalytic subunits and one 3'-5' exo/endoribonuclease subunit (**Figure 2.1A**, PDB 6H25 (4)). Three subunits make up the cap: EXOSC1, EXOSC2, and EXOSC3; while six subunits make up the hexameric core: EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, and EXOSC9. The catalytic ribonuclease DIS3 is located at the base of the complex. The cap and core assemble to form a channel through which RNA is threaded in a 3'-5' orientation to reach the catalytic subunit (59). Studies in various model systems have shown that the RNA exosome complex is essential for viability (1,9,44,50,60-63) and the complex is routinely referred to as ubiquitously expressed (64). To date, six of the nine structural subunit genes that make up the RNA exosome have been linked to conditions which each involve at least some degree of cerebellar atrophy (9,10,13,16,41,42,44,63,65). Why mutations in genes that encode structural subunits of the ubiquitously expressed RNA exosome complex give rise to neurological disease, which often impacts the cerebellum, is not well understood.

The RNA exosome complex plays critical roles in gene expression in subcellular compartments.

The RNA exosome localizes to both the nucleus and the cytoplasm and regulates many classes of RNA within these compartments. For example, in the nucleus, the RNA is critical for precise processing of ribosomal RNA to produce mature rRNA required for ribosomes (1,17). In addition, the nuclear RNA exosome targets RNA-DNA hybrids (R-loops), antisense RNAs, and small noncoding RNAs (snRNAs) for processing and/or degradation (18,25,56,66). In the cytoplasm, the RNA exosome targets mRNA for regulatory turnover, aberrant mRNAs for decay, such as mRNAs lacking a stop codon, and double stranded RNAs (dsRNAs) as a mechanism for viral defense (67-69).

The RNA exosome is hypothesized to target specific RNAs via interactions with proteins termed cofactors (33). In the nucleus, the RNA exosome associates with cofactors including the helicase MTR4, the TRAMP polyadenylation complex, and the MPP6 docking protein (5,70,71). Cytoplasmic RNA exosome cofactors include the rRNA channeling SKI complex (32). Given these critical interactions with the RNA exosome complex, any changes that alter the composition or conformation of the complex could have consequences for critical protein-protein interactions. Many important studies have provided insight into the function of the RNA exosome, including elegant structural studies (3-5,19,32,72,73), identification of key RNA substrates (2,18,24,56,67), and mechanistic insight into cofactor interactions (2,5,6,32,74). Many studies are performed using elegant biochemical approaches with reconstituted complexes, using genetic model organisms, or in cultured cells. These studies further unearth key questions about the requirements for RNA exosome function in specific cell types and tissues in multicellular organisms. These questions have been brought into sharp focus by recent studies linking mutations in genes encoding structural subunits of the RNA exosome to human diseases, which often have neurological involvement (44).

Clinical phenotypes of exosomopathies include neurological defects.

Missense mutations in genes encoding structural subunits of the RNA exosome give rise to a class of diseases termed exosomopathies. All exosomopathies described thus far include at least one missense *EXOSC* variant (63). Some patients are homozygous for the same missense variant, others are heterozygous for different missense variants, and some patients have a missense variant inherited in trans to a deletion (44). Additionally, exosomopathy patients have diverse clinical presentations including cerebellar atrophy, hypotonia, and respiratory difficulties (13). As the RNA exosome is required for many key cellular processes such as gene expression and translation, mutations that cause a complete loss of function of this complex are unlikely to be compatible with life. Thus, missense mutations linked to disease are expected to not be complete loss-of-function alleles but rather hypomorphic alleles.

As mutations in multiple genes encoding structural subunits of the RNA exosome have now been linked to disease (44), a formal postulation is that any pathogenic amino acid change that disrupts protein function could trigger a decrease in the level of that subunit and consequently the entire RNA exosome complex. Indeed, studies in fibroblasts from patients with *EXOSC* mutations support this hypothesis (10). However, if all pathological consequences were linked to loss of the RNA exosome complex, common pathology might be shared across patients with mutations in *EXOSC* genes. In contrast, a large number of distinct clinical phenotypes have been described for individuals with exosomopathies resulting from mutations in different subunit genes.

To illustrate the diversity of pathology described for exosomopathy patients, the following descriptions compile a number of the clinical diagnoses that have been identified in patients with mutations in *EXOSC* genes. One or more patients with a missense mutation in *EXOSC1* presented with hypoplastic cerebellum, cerebral atrophy, hyperextensibility of the skin, cardiomyopathy with reduced left ventricular ejection, and hypotonia (41,75). A single *EXOSC1* patient died due to renal

failure (41). One or more patients with a missense mutation in *EXOSC2* presented with borderline cerebellar hypoplasia, borderline intellectual disability, mild cortical and cerebellar atrophy, and hypothyroidism (42). Patients with missense mutations in *EXOSC3* have presented with symptoms of varying severity, classified as mild, moderate, or severe (45). The severity designations of *EXOSC3* patients are based on the genotype and clinical phenotype and clearly correlate (76). Some *EXOSC3* patients show severe hypotonia, progressive muscular atrophy, and postnatal and progressive cerebellar volume loss (45). Autopsies of *EXOSC3* patients demonstrated loss of neurons in the cerebellum, parts of the midbrain, and the anterior spinal cord (45). *EXOSC3* patients with mild PCH1 often survive into early puberty and reported respiratory failure in late stages of the disease, which was rarely the cause of death (45). *EXOSC3* patients with severe PCH1 had prenatal or congenital onset of cerebellar, pontine, and midbrain degeneration, as well as presented with severe hypotonia (45). These patients died in infancy from postnatal respiratory failure even under constant ventilation (45). In the case of *EXOSC5* mutations, patients required breathing support (43). One or more patients with *EXOSC5* mutations suffered from progressive hypotonia and respiratory impairment (43). MRIs of *EXOSC5* patients revealed reduced size of cerebellar vermis, brainstem, and pons (43). The echocardiogram of an *EXOSC5* patient showed anomalous coronary artery fistula (43). Several individuals with missense mutations in *EXOSC8* were reported to suffer from severe muscle weakness and died of respiratory failure before two years of age (9). MRIs of *EXOSC8* patients showed vermis hypoplasia and cortical atrophy (9). Autopsies of *EXOSC8* patients detected profound lack of myelin in cerebral, cerebellar white matter, and in the spinal cord (9). Patients with missense mutations in *EXOSC9* had progressively decreased strength, severe hypotonia, recurring pulmonary infections, and respiratory insufficiencies (10). MRIs revealed progressive cerebral and cerebellar atrophy. *EXOSC9* patients

showed rapid progressive muscle weakness and respiratory impairment combined with the presence of cerebellar atrophy and motor neuronopathy (10).

In contrast to mutations in genes encoding structural subunits of the RNA exosome, mutations in the *DIS3* gene, which encodes the catalytic ribonuclease (7), have been linked to multiple myeloma rather than diseases that largely affect the brain, like the exosomopathies previously described (77). *DIS3L*, an alternative ribonuclease that interacts with the nine structural subunits of the RNA exosome (59), has not been causatively linked to any inherited disease. Finally, three structural subunits of the RNA exosome, *EXOSC4*, *EXOSC6*, and *EXOSC7*, have not yet been linked to any reported pathology. However, we predict studies will soon surface describing new patients with mutations in these three structural subunits that may result in cerebellar atrophy.

With the small number of individuals diagnosed with exosomopathies thus far (44), more similarities may be revealed. Alternatively, disease pathology may be more tightly linked to the individual *EXOSC* genes altered. In this case, the location of the pathogenic amino acid change within a specific subunit may have a significant impact on protein-protein interactions, disrupting key interactions with specific cofactors or affecting the integrity of the complex. These changes may alter the function of individual subunits or the entire complex, ultimately affecting downstream RNA targets. There may be critical RNA targets that are important for the proper function of specific cells or tissues.

In this study, we employ publicly available transcriptomic data to explore expression of individual RNA exosome subunits in various human tissues. This analysis provides support for the characterization of the RNA exosome as ubiquitously expressed and reveals there is variability in the level of transcripts for individual RNA exosome in tissues linked to clinical pathology in

exosomopathies. However, the cerebellum has high levels of transcripts encoding virtually all RNA exosome subunits. This finding could suggest that cell types within the cerebellum require a high level of RNA exosome function for proper growth and/or maintenance.

Materials and Methods

Genotype-Tissue Expression (GTEx) project

GTEx release v8 includes whole genome sequencing (WGS) and RNA sequencing (RNA-seq) data from 54 tissues from 948 post-mortem individuals (312 females, 636 males; age 20-70). Each genotyped tissue has at least 70 samples. Violin plots and heatmap were generated to view median transcript per million (TPM) for the structural RNA exosome subunits: *EXOSC1*, *EXOSC2*, *EXOSC3*, *EXOSC4*, *EXOSC5*, *EXOSC6*, *EXOSC7*, *EXOSC8*, and *EXOSC9*. The number of genotyped donors for each tissue is as follows: brain/cerebellar hemisphere, n = 175; brain/cerebellum, n = 209; brain/cortex, n = 205; brain/frontal cortex, n = 175; brain/spinal cord (cervical c-1), n = 126; heart/atrial appendage, n = 372; heart/left ventricle, n = 386; kidney/cortex, n = 73; lung, n = 578; muscle/skeletal, n = 706; skin/not sun exposed (suprapubic), n = 517; thyroid, n = 574. The cerebellar hemisphere samples refer to the entire cerebellum and were preserved as frozen tissue. The cerebellum samples were procured from the right cerebellum and were preserved in a fixative. The cortex procured from the brain sampled the right cerebral pole cortex and was preserved in a fixative. The frontal cortex from the brain also sampled the right cerebral frontal pole cortex and was preserved as frozen tissue. The data used for the analyses described in this manuscript were obtained from the GTEx portal on April 20, 2023.

Results and Discussion

A comparative analysis of RNA exosome subunits reveals disparate requirements of transcripts in tissues.

Individuals with exosomopathies exhibit clinical phenotypes that extend to tissues beyond the brain. Tissues described in clinical diagnoses or linked with causes of death are chosen for this study. The varied clinical pathologies reported in RNA exosomopathies led us to include the following 12 human samples in this analysis: (1) cerebellar hemisphere/(2) cerebellum in the brain, (3) cortex/(4) frontal cortex in the brain, (5) spinal cord, (6) atrial appendage/(7) left ventricle in the heart, (8) kidney, (9) lung, (10) muscle, (11) skin, and (12) thyroid.

We examined the median expression as transcript per million (TPM) for each structural subunit of the RNA exosome in each of the tissues analyzed (**Figure 2.1B**). Violin plots show the levels of subunit transcripts compared between tissues. *EXOSC3* transcripts have the lowest overall TPM and *EXOSC5*, *EXOSC6*, and *EXOSC7* have the highest overall TPM across all tissues analyzed. Several of the RNA exosome subunit transcripts are expressed at the highest level in cerebellar hemisphere/cerebellum as compared to other tissues, including *EXOSC2*, *EXOSC6*, and *EXOSC9*. Importantly, the number of transcripts as reported by TPM values does not necessarily correlate with the amount of protein present in cells or tissues (78). *EXOSC2* and *EXOSC9* patients have distinct clinical diagnoses yet both of these subunits have some of the highest transcript levels (~20-25 median TPM) in cerebellar hemisphere/cerebellum.

To visualize the median TPM for clustered RNA exosome subunits and the tissues analyzed, we produced a heatmap using GTEx (**Figure 2.2A**). The cerebellar hemisphere/cerebellum, skin, and thyroid show the highest level of RNA exosome subunit transcripts. The atrial appendage and left ventricle of the heart have the lowest levels of subunit transcripts overall. *EXOSC3* transcripts are detected at the lowest levels across all tissues. To

compare all RNA exosome subunits for each tissue examined, we mapped the \log_{10} of the median TPM in violin plots (**Figure 2.2B**). Observations made from the heatmap are supported by the data provided in the violin plots.

We leveraged the GTEx consortium to visualize the transcript levels of RNA exosome subunits in brain regions, heart regions, kidney, lung, muscle, skin, and thyroid by median TPM in violin plots and heatmaps. These data show that overall, the tissues representing the cerebellum have high levels of RNA exosome subunit transcripts across the various structural subunits. Specifically, the cerebellar hemisphere/cerebellum has the highest levels of *EXOSC2*, *EXOSC6*, and *EXOSC9*. Skin has the highest levels of *EXOSC4*, *EXOSC5*, and *EXOSC7*. Thyroid has the highest levels of *EXOSC1*, *EXOSC3*, and *EXOSC8*. The left ventricle of the heart has the lowest transcript levels of *EXOSC1*, *EXOSC2*, *EXOSC3*, *EXOSC6*, and *EXOSC9*. The brain cortex has the lowest transcript levels of *EXOSC7* and *EXOSC8*. Muscle has the lowest levels of *EXOSC4*, and kidney has the lowest levels of *EXOSC5*. Additionally, these data show that *EXOSC1* has the highest transcript levels in the lung and thyroid. *EXOSC5* has the highest transcript levels in the cortex, spinal cord, and muscle. *EXOSC6* has the highest transcript levels in the cerebellar hemisphere/cerebellum and frontal cortex. *EXOSC7* has the highest transcript levels in the atrial appendage and left ventricle of the heart, kidney, and skin. *EXOSC3* shows the lowest transcript level in all tissues examined except the kidney; however, levels in the kidney are relatively low compared with other subunit transcripts (*EXOSC1*, *EXOSC4*, *EXOSC5*, *EXOSC6*, *EXOSC7*, and *EXOSC8*). Levels of *EXOSC2* transcript are lowest in kidney.

Across nearly all tissues examined, *EXOSC3* transcript levels are detected at the lowest levels. *EXOSC3* levels in the kidney are only slightly above the lowest (lowest TPM = 5.29, highest TPM = 19.42, *EXOSC3* TPM = 5.53). Although little is known about the assembly of the RNA

exosome complex *in vivo*, there could be a limiting subunit that determines overall levels of the complex. The observation that the *EXOSC3* transcript is low in most tissues examined could suggest that *EXOSC3* levels are limiting for RNA exosome complex assembly or steady-state levels. Interestingly, mutations in the *EXOSC3* gene were the first identified and linked to human disease (79). Perhaps, if levels of *EXOSC3* are limiting, even small changes that alter the function or level of *EXOSC3* could cause pathology. Indeed, the most mutations have now been identified and reported in *EXOSC3* (76), which could be consistent with this RNA exosome subunit being most vulnerable to even minor changes in function or levels. Alternatively, as *EXOSC3* was the first subunit of the RNA exosome linked to disease, the larger number of cases could simply represent ascertainment bias. As additional exosomopathy cases are identified and described, differences in the numbers of patients and types of pathologies associated with mutations in different *EXOSC* genes will likely be clarified.

The results reported here suggest that some tissues, such as the cerebellar hemisphere/cerebellum, skin, and thyroid may require more RNA exosome function as compared to other tissues such as the heart, brain cortex, muscle, and kidney. The data presented here may explain why exosomopathy patients with missense mutations in structural RNA exosome subunit genes have the most well-defined phenotypes in the cerebellum as compared to other tissues. However, steady-state transcript levels often do not directly translate to protein levels (78) as there are many additional regulatory steps that determine steady-state protein levels. Further proteomic analysis of RNA exosome subunits in specific cell types, such as those that are abundant in the cerebellum, could provide insight into why mutations in genes encoding structural subunits of the RNA exosome often cause cerebellar pathology. RNA exosome function is determined by many protein-protein interactions with cofactors. The cell-specificity and/or cell-specific interactions of

well-defined RNA exosome cofactors has not yet been determined and novel interactors are continuously discovered (66). Changes in cell-specific interactions could explain the phenotypes observed in exosomopathy patients.

This study provides insight into the expression of various structural subunits of the RNA exosome in human tissues, supporting the common statement that the RNA exosome complex is ubiquitously expressed. This work reveals high transcript levels for multiple RNA exosome subunits in cerebellar samples, which could begin to explain why exosomopathies often present with cerebellar involvement.

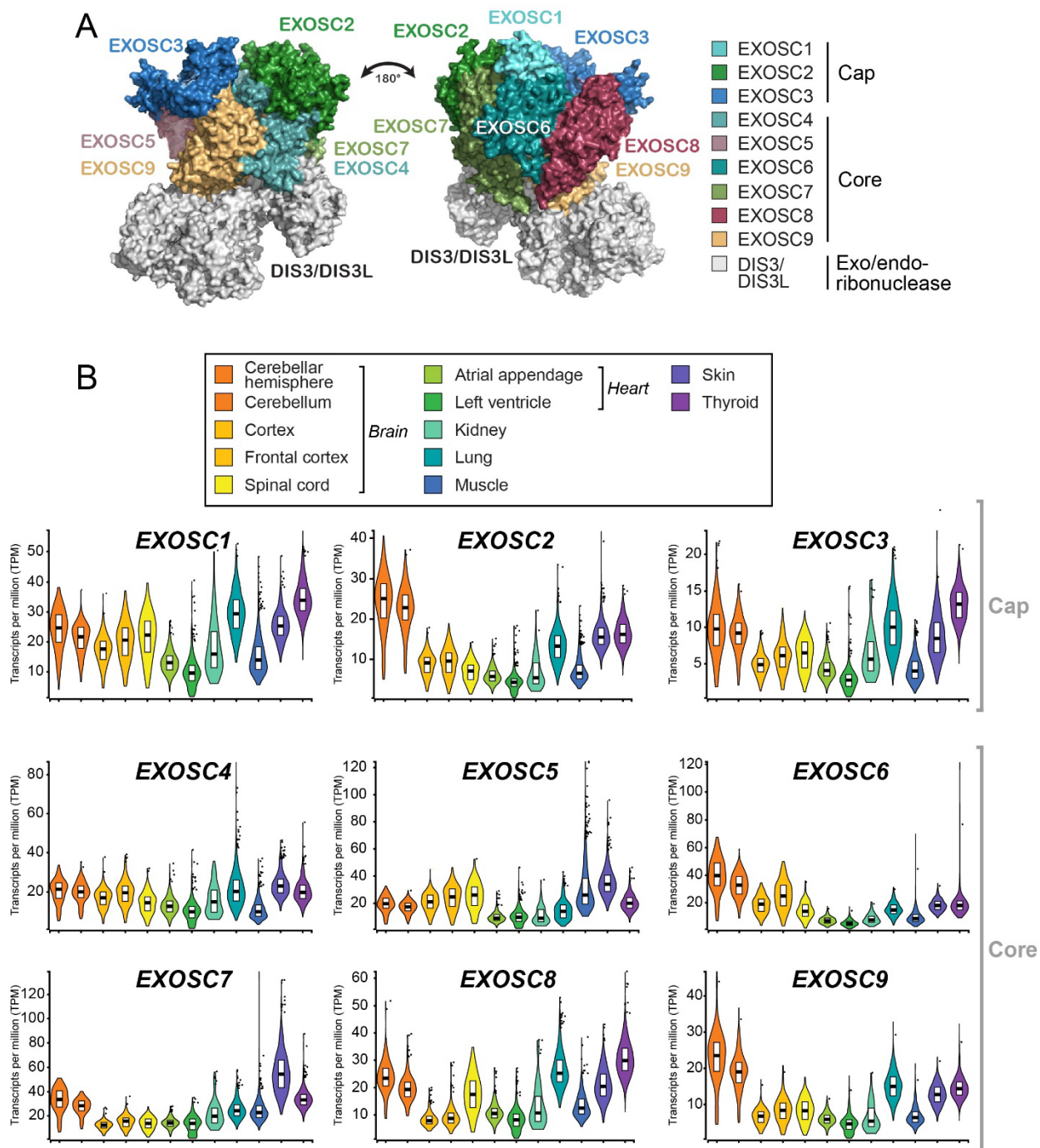


Figure 2.1. The RNA exosome subunit transcripts are ubiquitously expressed in human tissue.

(A) The RNA exosome is a ten-subunit complex that targets multiple classes of RNA. Nine of the ten subunits of the RNA exosome are structural and nonenzymatic. Three subunits compose the cap: EXOSC1, EXOSC2, EXOSC3, and six subunits compose the hexameric core: EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, EXOSC9. The base subunit, DIS3 or DIS3L, has catalytic ribonucleolytic activity. The image is adapted from the PDB structure 6H25 (4) . (B) The nine structural subunits of the RNA exosome are expressed in all tissues examined. Transcript levels of each subunit vary; however, several subunits show high levels in the cerebellar hemisphere/cerebellum. Each structural subunit is analyzed in tissues that have been previously reported in the clinical diagnoses or causes of death for exosomopathy patients. As described in detail in the Materials and Methods, tissues examined for each RNA exosome subunit include the brain (cerebellar hemisphere, cerebellum, cortex, frontal cortex, and spinal cord), heart (atrial appendage and left ventricle), kidney, lung, muscle, skin, and thyroid. RNA levels of subunits are presented in violin plots by median transcript per million (TPM) and vary from tissue to tissue.

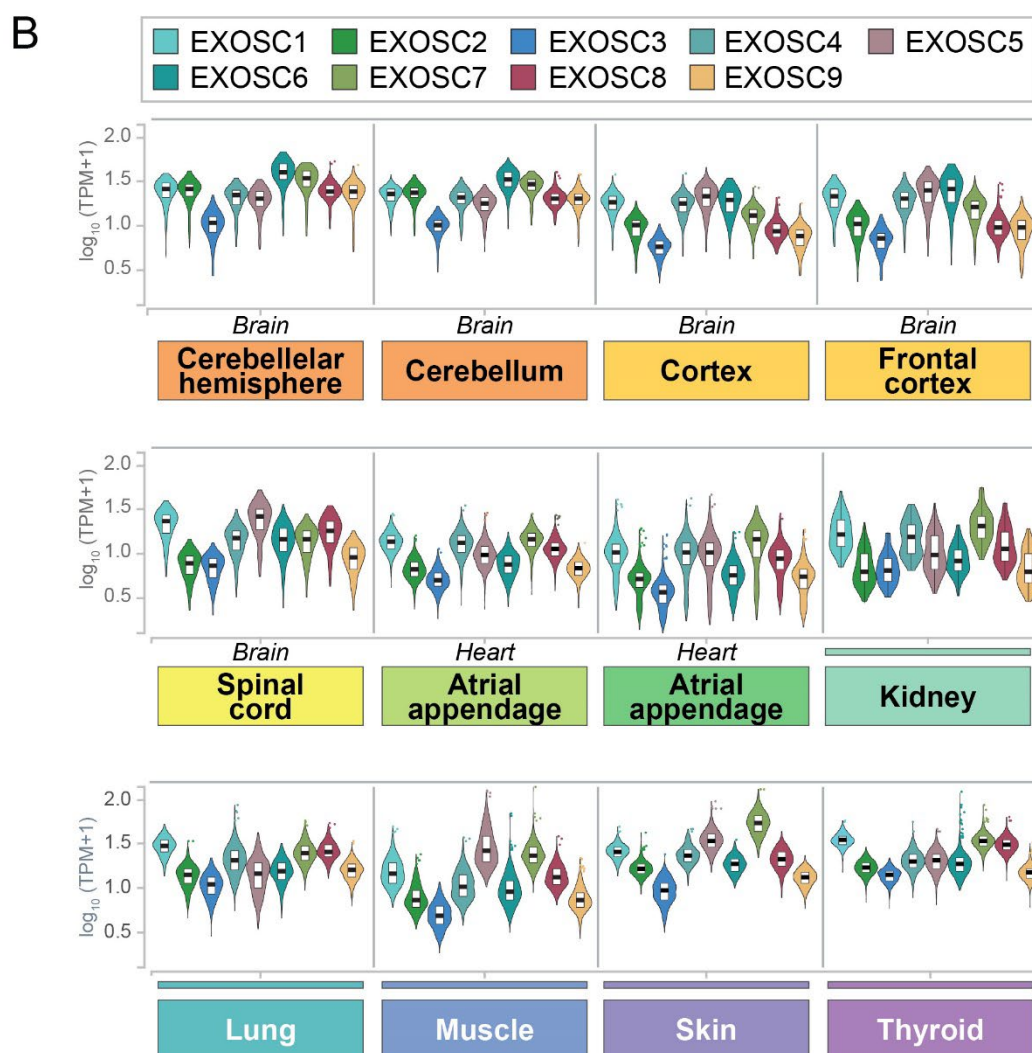
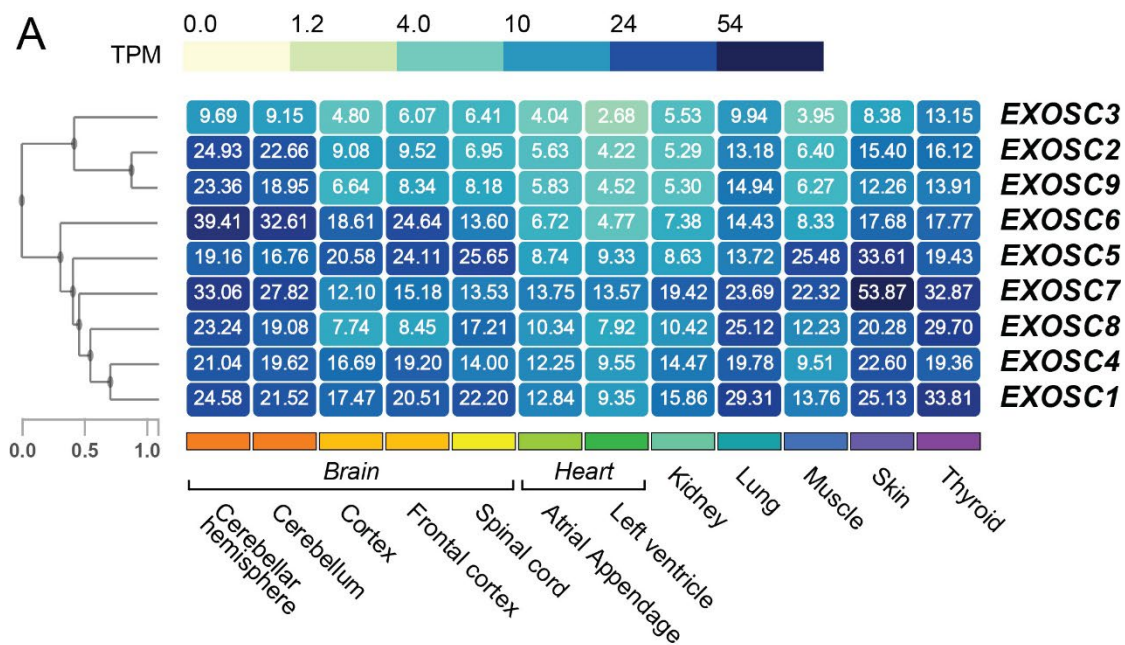


Figure 2.2. RNA exosome subunit transcript levels are high in the cerebellar hemisphere/cerebellum compared with other tissues.

(A) RNA exosome subunit transcript levels across a set of tissues implicated in clinical diagnoses or causes of death in exosomopathy patients are shown in a heatmap. The data generated provide visualization of transcript levels across tissues as illustrated by median TPM and are clustered by RNA exosome subunit. (B) A comparative analysis of subunits reveals variable transcript levels across different tissues. Median TPM is adjusted by \log_{10} TPM for comparison.

Chapter 3

RNA helicase DDX1 interacts with the RNA exosome to modulate R-loops.

The following chapter was submitted to the *Journal of Biological Chemistry* and is currently in revision. The preprint is available:

de Amorim, J. L., Leung, S. W., Haji-Seyed-Javadi, R., Hou, Y., Yu, D. S., Ghalei, H., ... & Corbett, A. H. (2023). The RNA helicase DDX1 associates with the nuclear RNA exosome and modulates R-loops. *bioRxiv*, 2023-04.

Abstract

The RNA exosome is a ribonuclease complex that mediates both RNA processing and degradation. This complex is evolutionarily conserved, ubiquitously expressed, and required for fundamental cellular functions, including rRNA processing. The RNA exosome plays roles in regulating gene expression and protecting the genome, including modulating the accumulation of RNA-DNA hybrids (R-loops). The function of the RNA exosome is facilitated by cofactors, such as the RNA helicase MTR4, which binds/remodels RNAs. Recently, missense mutations in RNA exosome subunit genes have been linked to neurological diseases. One possibility to explain why missense mutations in genes encoding RNA exosome subunits lead to neurological diseases is that the complex may interact with cell- or tissue-specific cofactors that are impacted by these changes. To begin addressing this question, we performed immunoprecipitation of the RNA exosome subunit, EXOSC3, in a murine neuronal cell line (N2A) followed by proteomic analyses to identify novel interactors. We identified the putative RNA helicase, DDX1, as an interactor. DDX1 plays roles in double-strand break repair, rRNA processing, and R-loop modulation. To explore the functional connections between EXOSC3 and DDX1, we examined the interaction following double-strand breaks, and analyzed changes in R-loops in N2A cells depleted for EXOSC3 or DDX1 by DNA/RNA immunoprecipitation followed by sequencing (DRIP-Seq). We find that EXOSC3 interaction with DDX1 is decreased in the presence of DNA damage and that loss of EXOSC3 or DDX1 alters R-loops. These results suggest EXOSC3 and DDX1 interact during events of cellular homeostasis and potentially suppress unscrupulous expression of genes promoting neuronal projection.

Introduction

The RNA exosome is a 10-subunit ribonuclease complex responsible for processing and degradation of many classes of RNA in all eukaryotes and many archaea. The ribonuclease activity of the RNA exosome is critical for both RNA quality control and precise processing of key RNAs, including ribosomal RNA (rRNA) (1,17,18). As illustrated in **Figure 3.1A**, the 10 subunits of the RNA exosome are organized into a non-catalytic cap, composed of three subunits (EXOSC1, EXOSC2, and EXOSC3), a non-catalytic core ring, comprising six subunits (EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, and EXOSC9), and one catalytic 3'-5' exo/endoribonuclease subunit (DIS3 or DIS3L) (2-6,80). Most target RNAs are threaded through the cap and central channel of the RNA exosome to reach DIS3 or DIS3L for processing and/or degradation (3,7,21,81). Studies in yeast and other model systems have shown that the RNA exosome complex is essential (1,9,43,50,60-62) and ubiquitously expressed (64). Although this complex has been studied for decades, key questions, such as how the RNA exosome is targeted to specific RNAs, remain to be answered.

The RNA exosome complex processes/degrades multiple classes of RNAs in both the nucleus and the cytoplasm. The best-defined role of the RNA exosome is the processing of precursor rRNA to mature rRNA for the production of ribosomes (1,17,35,82,83). Other RNAs in the nucleus/nucleolus targeted by the RNA exosome include RNA from RNA/DNA hybrids, commonly known as R-loops, promoter upstream transcripts (PROMPTs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and other non-coding RNAs (ncRNAs), such as transcription start site (TSS)-associated antisense transcripts (xTSS-RNA) (18,24,25,29,30,56,67). In the cytoplasm, the RNA exosome targets aberrant transcripts for degradation, mRNAs for regulatory turnover, and viral RNAs as a cellular immune response (67-69,84-86). The RNA

exosome has also been implicated in DNA double-strand break repair by homologous recombination (HR), potentially targeting aberrant transcripts produced upon DNA damage (87). The transcripts that are targeted in response to DNA damage are often within R-loop structures, which occur naturally during transcription. An accumulation of R-loops may have deleterious effects, leading to double-strand breaks and genomic instability (88). Studies suggest that the RNA exosome is poised to degrade released RNA after R-loops are unwound by RNA/DNA helicases (89).

The RNA exosome plays a critical role in cells by degrading and/or processing many transcripts in different cellular compartments. Thus, missense mutations in genes encoding structural subunits of the complex are linked to several human diseases, termed exosomopathies (44,63). The first link between the RNA exosome complex and disease described a patient with a missense mutation in *EXOSC3*, which causes the neurological disease pontocerebellar hypoplasia type 1B (PCH1B) (16). This autosomal recessive disease is characterized by severe atrophy and progressive hypoplasia of the pons and cerebellum (45,90,91). Since the initial report of mutations in *EXOSC3*, more mutations have been identified and described in *EXOSC1*, *EXOSC2*, *EXOSC5*, *EXOSC8*, and *EXOSC9* (9,10,41-43,92). All patients, with the exception of individuals with mutations in *EXOSC2*, suffer from cerebellar atrophy, at least to some extent. Patients with mutations in *EXOSC2* present with a syndromic condition that consists of short stature, hearing loss, retinitis pigmentosa, distinctive facies, and mild intellectual disability (15,42). Why mutations in genes encoding structural subunits of the RNA exosome impact the cerebellum is not at all clear.

A number of protein cofactors associate with the RNA exosome to confer RNA target specificity. Several RNA exosome cofactors that were originally identified and characterized in

budding yeast are conserved in human (2,13,27,32,37,93,94). For example, the RNA exosome requires helicases for proper RNA processing as RNAs with significant secondary structure are unable to enter the central channel of the hexameric ring of the RNA exosome (95). A nuclear helicase, termed MTR4 (alternatively named SKIV2L2), interacts with the RNA exosome and facilitates processing of RNA in the nucleus (5,6). A well characterized cytoplasmic helicase, termed SKIV2L in humans or Ski2 in budding yeast, directly interacts with the RNA exosome and assists in cytoplasmic rRNA processing (32,37). The RNA exosome also interacts with nuclear scaffolding proteins, such as MPP6 (alternatively named M-phase phosphoprotein 6), and other associated ribonucleases, such as EXOSC10 (human) or Rrp6 (budding yeast) (5,6). One potential hypothesis to explain how single amino acid changes in structural subunits of the RNA exosome cause disease is that modest changes in the subunits alter interactions of the complex with cofactors required to target and subsequently process or degrade specific RNAs. As the RNA exosome is essential for fundamental processes, such as the production of mature ribosomes, a complete loss of function in patients seems unlikely. Why the majority of missense mutations identified in patients with exosomopathies cause clinical consequences most notable within regions of the brain remains unclear.

In this study, we examined the interactome of the RNA exosome in a neuronal cell line. An unbiased mass spectrometry approach identified a number of candidate binding partners. We identified the putative RNA helicase DDX1 as a protein that interacts with EXOSC3 and explored the shared functions. We found that DDX1 interacts with EXOSC3 in the nucleus, that the interaction is DNA damage-sensitive, and that the depletion of EXOSC3 or DDX1 results in significant changes in R-loops. Together, these findings suggest a novel aspect of RNA exosome function and regulation that is required for gene expression control.

Results

Proteomics reveal a suite of EXOSC3 interactors.

To identify RNA exosome-interacting proteins in a neuronal cell line, we transiently transfected a plasmid encoding myc-tagged EXOSC3 into a murine neuroblastoma cell line (N2A) and purified co-precipitated proteins using anti-myc magnetic beads as described in Materials and Methods. Immunoblots shown in **Figure 3.1B** confirm that myc-EXOSC3 is enriched in the Bound fraction compared with the Vector control. The core subunit EXOSC9 of the RNA exosome co-precipitates with myc-EXOSC3, suggesting myc-tagged EXOSC3 associates with the RNA exosome complex. We then subjected the immunoprecipitates to LC-MS/MS as described in Materials and Methods. The table in **Figure 3.1C** shows that all subunits of the RNA exosome complex were detected in the myc-EXOSC3 immunoprecipitation as compared to samples from the control.

The RNA exosome-interacting proteins identified by mass spectrometry were analyzed using the Panther Gene Ontology (GO) program and organized by protein class (**Figure 3.2A**). We excluded all proteins for which the peptide spectra matches (PSM) \log_2 ratio was less than or equal to zero. We examined 955 proteins for which the PSM equated to greater than zero. The largest category of the GO protein classes is “translation” containing 120 proteins and the second largest is “nucleic acid metabolism and binding” with 114 proteins. Within the latter category, all the RNA exosome subunits and some known cofactors, including MTR4 (also known as SKIV2L2) and MPP6, are present. **Figure 3.2B** shows cofactors (green) and potential RNA exosome-interacting candidates selected for further analysis (blue). The PSM and peptide numbers are low for even well-established cofactors, and therefore we used literary analysis and a higher

PSM and peptide number to inform decisions to explore specific interactors. For this analysis, we opted to focus on Nucleic acid metabolism/binding instead of Translation because several cytoplasmic interactions between the RNA exosome and cofactors have been well characterized (37,93) and interactions between the RNA exosome and ribosome subunits have been defined (32,94).

Putative helicase DDX1 interacts with the RNA exosome in the nucleus.

To investigate potential endogenous protein candidates, we first produced a custom antibody targeting EXOSC3 and tested antibody specificity using two independent *EXOSC3* siRNAs (**Figure 3.S1A**). We additionally performed cell fractionation to isolate the nucleus from the cytoplasm in a neuronal cell line, drawing from evidence of known compartment-specific RNA exosome interactions, including the nuclear MTR4 and the cytoplasmic SKI complex (**Figure 3.S1B**) (33,59). While we analyzed a number of candidates, we focused on the putative RNA helicase DDX1 for several reasons. RNA helicases play critical roles in various aspects of RNA metabolism, including RNA degradation and processing (96). Additionally, multiple conserved RNA exosome cofactors are helicases (5,32). Based on sequence homology, DDX1 is a putative helicase containing a conserved DEAD amino acid sequence motif shared by nucleic acid-unwinding, ATP-binding, DEAD-box proteins (96-98). DDX1 differs from other members of the DEAD-box family, as it includes an N-terminal SPRY protein interacting domain upstream of two helicase domains, between the phosphate-binding P-loop and the single-strand DNA binding Ia motifs (**Figure 3.3A**) (96). The DDX1 protein is implicated in rRNA processing (99), R-loop formation (100), and double-strand break repair (101,102).

Initial studies employed epitope-tagged, transiently transfected EXOSC3. To test interactions with endogenous EXOSC3, we raised a rabbit polyclonal antibody against mouse

EXOSC3. To test the specificity of this newly generated antibody, we depleted cells of EXOSC3 by transfecting N2A cells with two independent siRNA oligonucleotides that target *EXOSC3* and performed immunoblotting (**Figure 3.S1A**). In total N2A cell lysate, the antibody detects a prominent band at the predicted size of EXOSC3 (calculated molecular weight of 29.5 kDa). This band is significantly decreased when cells are treated with either siRNA targeting *EXOSC3* as compared to the Scramble control, providing evidence for the specificity of the antibody generated.

Initial attempts to validate putative RNA exosome interacting proteins identified by mass spectrometry in whole cell lysate were unsuccessful. Thus, we considered the fact that many RNA exosome cofactors localize to specific cellular compartments (33), and we examined interactions with endogenous EXOSC3 using cellular fractionation as described in Materials and Methods. Fractionation was confirmed via immunoblotting for the cytoplasmic marker HSP90 (103) and the nuclear protein B23 (104) (**Figure 3.S1B**). A band corresponding to the B23 is enriched in the nuclear fraction and absent in the cytoplasmic fraction, indicating efficient nuclear isolation. However, a band corresponding to HSP90 in the nuclear fraction suggests some cytoplasmic adulteration. We immunoprecipitated endogenous EXOSC3 from both the nuclear and cytoplasmic fractions, then used SDS-PAGE and immunoblotting to assess co-purification. DDX1 is detected in the Input of both the nuclear and cytoplasmic lysates, consistent with reported localization (**Figure 3.3B**) (105,106). However, DDX1 is present in the Bound fraction for only the Nucleus and not the Cytoplasm. DDX1 was not detected in any of the Bound fractions for Ctrl IgG samples. EXOSC9, a core component of the RNA exosome complex is detected in both the Cytoplasm and Nucleus Inputs as well as the Bound fractions, consistent with the fact that the RNA exosome complex is present in both compartments. EXOSC3 is enriched in the Bound

fractions in both the Cytoplasm and Nucleus, and not in control IgG (Ctrl IgG). These data suggest a compartment-specific interaction between EXOSC3 and DDX1 in the nuclear fraction.

To assess whether the interaction detected between EXOSC3 and DDX1 is RNA- or DNA-dependent, we treated the nuclear N2A cell lysate with several enzymes including RNase A, RNase T1, RNase H, DNase I, and benzonase. A urea-PAGE gel confirmed that RNase A, RNase T1, and benzonase degraded total RNA from the lysate (data not shown). We found that the treatment with RNase A or benzonase significantly increases the interaction between EXOSC3 and DDX1 with no detectable effect on the steady-state level of either protein (**Figure 3.3C**). This experiment suggests the interaction is not RNA-dependent and the significant increase in the interaction with RNase A treatment has been suggested to indicate a protein-protein interaction (107). These results could suggest that RNA in fact interferes with the interaction between EXOSC3 and DDX1.

The interaction between EXOSC3 and DDX1 decreases in response to DNA damage.

The interaction between EXOSC3 and DDX1 is nuclear-specific and is increased in the absence of RNA. We sought to further characterize the interaction between the two proteins. Because both the RNA exosome and DDX1 play roles in DNA damage repair (87,101,102), we tested whether the interaction between the proteins is sensitive upon induction of DNA double-strand breaks. We treated N2A cells with 5 μ M camptothecin (CPT), a topoisomerase inhibitor that induces double-strand breaks (108), or control phosphate-buffered saline (PBS) at 37°C for one hour. To confirm DNA damage, we used immunofluorescence to detect γ H2AX, a classic marker of double-strand breaks (109). As shown in **Figure 3.4A**, after treatment with camptothecin, the γ H2AX signal increases markedly compared with the PBS control with no detectable relocalization of EXOSC3 or DDX1. We used the EXOSC3 antibody to immunoprecipitate endogenous EXOSC3 from the nuclear fraction of cells treated with CPT and

probed for DDX1 (**Figure 3.4B**). The interaction between EXOSC3 and DDX1 is significantly reduced following treatment with camptothecin (**Figure 3.4C**). In contrast, there is no change in the interaction between EXOSC3 and EXOSC9. In concordance with Figure 5A, the compartmental localization of the interaction between EXOSC3 and DDX1 does not change upon damage (**Figure 3.4B**). Thus, the interaction between EXOSC3 and DDX1 is substantially decreased in response to the induction of double-strand breaks. These results suggest that the roles EXOSC3 and DDX1 currently play in DNA damage repair are distinct from one another.

Depletion of EXOSC3 or DDX1 results in rRNA processing defects.

To explore potential shared functions of the RNA exosome and DDX1, we optimized conditions to siRNA deplete each protein from N2A cells. We transfected N2A cells with siRNA scramble control (Scramble) or EXOSC3 siRNA, then performed an immunoblot to confirm depletion (**Figure 3.5A**). Knockdown was quantified across three biological replicates (**Figure 3.5B**). Similarly, N2A cells were transfected with DDX1 siRNA, and depletion was confirmed by immunoblot (**Figure 3.5C**). The knockdown was quantified across the three biological replicates (**Figure 3.5D**). Thus, we were able to substantially deplete each protein to less than 15% remaining (**Figure 3.5B, 3.5D**), providing a model to explore and compare the consequences of loss of each of these proteins.

The RNA exosome has a well-defined role in rRNA processing and maturation (1,17,35,83,110). DEAD-box helicases such as DDX1 also play a critical role in RNA processing and genome stability (107,111). A previous study that analyzed DDX1 knockout mouse embryonic stem cells (mESCs) employed a pulse-chase experiment utilizing [³H]-uridine-labeled samples and showed an accumulation of precursor 28S rRNA and mature 18S rRNA, suggesting a role for DDX1 in rRNA processing (112). We performed a detailed analysis to explore rRNA processing

in cells depleted of either EXOSC3 or DDX1. RNA quality was measured by agarose gel (**Figure 3.S2A**) and High Sensitivity ScreenTape assay (**Figure 3.S2B**). Using northern blotting to detect specific ribosomal RNA precursors, we examined which rRNA species are affected upon loss of either EXOSC3 or DDX1. **Figure 6A** depicts the steps of murine rRNA processing from the early precursor 47S to mature rRNA. Ribosomal RNA processing begins with the 47S precursor, which generates several downstream precursors, including 32S and 12S (113,114). The 47S precursor also produces the 18S, 5.8S, and 28S mature rRNAs, as well as the internal transcribed spacers 1 and 2 (ITS1 and ITS2), and the 5' and 3' external transcribed spacers (5'ETS and 3'ETS). To capture these precursors, we used probes specific to the ITS2 sequence. **Figure 6B** shows that the steady-state levels of the 41S, 20S, and 5.8S precursors decrease when EXOSC3 is depleted compared to the Scramble control. The levels of the 29S precursor decrease upon EXOSC3 depletion and inversely increase upon DDX1 depletion. 12S precursor increases when EXOSC3 is depleted compared to the Scramble control, consistent with the most well-defined role of the RNA exosome in 3' trimming to produce mature 5.8S (1,17,68,83,110). In contrast, when cells are depleted of DDX1, there is a decrease in the steady-state levels of the 12S precursor compared to the Scramble control. When using probes to detect 5.8S rRNA in the same samples, we detect a decrease in 5.8S level after depletion of EXOSC3, but no significant change when DDX1 is depleted. The 7SL signal recognition particle (SRP) transcript, which is not a target of the RNA exosome (115), is used as a loading control. The northern blot data are quantified for all analyses in **Figure 6C** and normalized to the loading control. The data from **Figure 8B** and **Figure 8C** are also summarized within **Figure 8A**, as indicated by the up- and down-arrows to denote statistically significant increases or decreases in these RNA species. The precursors are targeted using specific

probes detailed in **Table 3.S1**. Although both EXOSC3 and DDX1 clearly have an impact on rRNA processing or maturation, the roles in this process appear to be independent of each other.

R-loops are globally reduced upon depletion of EXOSC3 or DDX1.

Multiple studies have linked DDX1 to R-loops, which are three-strand nucleic acid structures comprised of an RNA-DNA hybrid and a single strand of DNA (100,111,116). DDX1 has been reported to co-precipitate with R-loops and promote R-loop formation by unwinding complex RNA (100,111,116). Studies have also linked the RNA exosome to R-loop regulation in murine B-cells (117). To understand the impact EXOSC3 and DDX1 may have on R-loops, we performed DNA/RNA-immunoprecipitation followed by high-throughput sequencing (DRIP-seq) in cells depleted of either EXOSC3 or DDX1. In cells depleted of EXOSC3, we identified 722 significantly increased R-loop regions and 935 decreased R-loop regions (**Figure 3.7A**, left, $n = 3$, $FDR < 0.05$). In cells depleted of DDX1, 638 increased R-loop regions and 1,058 decreased R-loop regions are identified (**Figure 3.7A**, right, $n = 3$, $FDR < 0.05$). We then compared the R-loop regions that showed statistically significant changes in cells depleted of EXOSC3 or DDX1. We found that 140 out of 722 increased R-loop regions (~19%) in EXOSC3-depleted cells are also increased in DDX1-depleted cells (**Figure 3.7B**, left, Chi-squared test, p -value < 0.0005) and 425 out of 935 decreased R-loop regions (~45%) in EXOSC3 knockdown cells are also decreased in DDX1 knockdown cells (**Figure 3.7B**, right, Chi-squared test, p -value < 0.0005). These results suggest that the two proteins may cooperate to alter a common set of R-loops.

EXOSC3- and DDX1-depleted cells share common R-loop regions that increased or decreased. We employed Panther Gene Ontology (GO) analysis on these R-loops (**Figure 3.7C**). The R-loop regions that are increased in both EXOSC3 and DDX1 knockdowns ($n = 140$) are enriched in the categories of protein folding, histone modifications, stress responses, and RNA

metabolic processes when grouped by fold enrichment. The R-loop regions that are decreased in both EXOSC3 and DDX1 knockdown conditions ($n = 425$) are also enriched in the RNA metabolism category, but additionally in categories including axon guidance and neuronal development.

In **Figure 3.7D**, we grouped both increased and decreased R-loop regions by RNA transcript class present within the region corresponding to altered R-loops. Both EXOSC3 and DDX1 depletion affects R-loops within genes encoding different classes of RNA. We grouped all changed R-loop regions, both increased and decreased, after EXOSC3 depletion and found that the largest category of RNA significantly affected is protein-coding ($n = 1,496$). We then excluded protein-coding genes from the analysis to allow a clearer view of the non-coding transcripts and identified 80 ncRNAs, 55 microRNAs, 9 snoRNAs, 8 pseudoRNAs, 6 lncRNAs, 1 scRNA, and 1 telomerase RNA within the altered R-loop regions. An analysis following DDX1 depletion revealed similar results. We excluded 1,509 R-loop regions that mapped to protein-coding genes and identified 95 ncRNAs, 61 microRNAs, 13 snoRNAs, 12 pseudoRNAs, and 5 lncRNAs. Noncoding RNAs (ncRNAs) are those transcripts that do not currently have a more specific distinction.

In parallel with DRIP-seq, we employed RNA sequencing (RNA-seq) after ribosomal RNA depletion to identify transcripts altered by depletion of EXOSC3 or DDX1. The pipeline employed for this analysis focused on coding regions, so data presented represent changes in mRNA transcripts. We identified 1,757 significantly increased transcripts and 2,192 decreased transcripts in EXOSC3 knockdown cells (**Figure 3.8A**, left, $n = 3$, $FDR < 0.05$). In DDX1 knockdown cells, 734 increased transcripts and 968 decreased transcripts are identified (**Figure 3.8A**, right, $n = 3$, $FDR < 0.05$). We then compared the transcripts that showed statistically significant changes in

EXOSC3- and DDX1-depleted cells. We found that 322 out of 1,757 increased mRNA transcripts (~18%) in siEXOSC3 cells are also increased in siDDX1 cells (**Figure 3.8B**, left, Chi-squared test, p-value < 0.0005) and 599 out of 2,192 decreased mRNA transcripts (~27%) in siEXOSC3 cells are also decreased in siDDX1 cells (**Figure 3.8B**, right, Chi-squared test, p-value < 0.0005), suggesting a coordination between EXOSC3 and DDX1 in commonly regulating a critical set of genes.

With both DRIP-seq and RNA-seq datasets in hand, we created a pipeline to compare results and filter the DRIP regions through the differential mRNA sequencing data to focus on the overlapping similarity between R-loops and changes in transcript level at those specific R-loop regions (**Figure 3.9A**). We applied the pipeline to generate a heatmap to illustrate the DRIP regions and mRNA transcripts that changed upon depletion of either EXOSC3 or DDX1 (**Figure 3.9B**). There are 466 R-loop regions that overlap in the RNA-sequencing data for both depletions. Of these overlapping R-loop regions, 103 are increased, and 363 regions are decreased. Many of the decreased regions mapped to genes that are involved in RNA metabolism, RNA regulation, translation processes, and neuronal development. We employed the Integrative Genomics Viewer (IGV), which enables the visualization of these regions (**Figure 3.9C**). To illustrate the effect on specific loci, we focused on two genes that contained R-loop regions significantly changed within transcripts that are also significantly changed, and which fell under the biological processes aforementioned: *Ints6* and *Celf4*. The red marker above the gene (orange) in the IGVs of *Ints6* denotes the R-loop region, indicating a potential reduction in gene expression at these loci. In the three panels on the right, *Celf4* R-loop regions are marked with a red line above the gene (orange). The two left IGV panels corresponding to *Celf4* show increased R-loop and transcript regions. In the far right IGV panel corresponding to *Celf4*, the R-loop regions are decreased, and transcript

regions are increased. The changes in R-loop regions at these loci indicate potential changes in gene regulation and expression upon the depletion of EXOSC3 or DDX1. All IGV panels show regions that have statistically significant changes compared to Scramble. Quantification of *Ints6* and *Celf4* transcripts show statistical significance (**Figure 3.9D**). Altogether, these data indicate that depletion of either EXOSC3 or DDX1 results in changes in R-loop regions that do not necessarily correspond to a similar change in transcript levels for that same gene.

Discussion

In this study, we used a proteomic approach to identify RNA exosome-associated proteins in a neuronal cell line and identified an interaction between EXOSC3 and the putative RNA helicase DDX1. Although each protein is present in both the nucleus and the cytoplasm, the interaction was only detected in the nuclear fraction. To explore possible shared functions between EXOSC3 (RNA exosome) and DDX1 in which each of these proteins are implicated, we examined the EXOSC3-DDX1 interaction upon DNA damage, and effects on rRNA processing and on R-loop distribution upon depletion of EXOSC3 or DDX1. Our results suggest that EXOSC3 and DDX1 participate in shared regulation of R-loops and the transcripts produced within the genomic regions that form those R-loops. Taken together, these data define a potential mechanism by which an interaction between the RNA exosome complex and an RNA helicase could modulate R-loops.

The RNA exosome complex interacts with a number of proteins, termed cofactors, to confer specificity for different RNA targets (27). Many studies, including a number of elegant structures (3-7,19,36,37,74), show that the RNA exosome consists of a core set of subunits and cofactors that are present in one-to-one stoichiometry. The core RNA exosome then interacts with a number of different proteins to facilitate degradation or processing of many different RNA

targets. These protein-protein binding events may be transient as the RNA exosome interacts with multiple cofactors at shared binding sites that have been revealed by biochemical and structural studies (5). Indeed, only a single peptide was identified in the mass spectrometry data in this study for the well-characterized RNA exosome cofactors, MTR4 (alternatively named SKIV2L2) and MPP6. These results provide evidence that the interactions between the RNA exosome and cofactors are likely to be dynamic and transient.

Beyond the previously characterized RNA exosome cofactors, this study identified several candidate RNA exosome-interacting proteins, which are located in different cellular compartments. For example, DDX1 is present in both the nucleus and the cytoplasm (106), while the Pumilio proteins, PUM1 and PUM2, are reported to control RNA stability exclusively in the cytoplasm (118). Thus, further studies could explore whether the interactions identified in whole cell lysate occur preferentially in one cellular compartment or another. Although DDX1 is readily detected in both the nucleus and the cytoplasm (**Figure 3.3B**) (106), robust interaction between EXOSC3 and DDX1 was only detected in the nuclear lysate. This compartment-specific interaction could mean that the RNA exosome and DDX1 interact in the context of chromatin. However, treatment with DNase I did not substantially decrease the EXOSC3-DDX1 association (**Figure 3.3C**). One possibility as to why the EXOSC3-DDX1 interaction is compartment-specific is that post-translational modifications (PTMs) could modulate binding. Studies in *S. pombe* have revealed PTMs in Dis3, Mtr4, Rrp40 (EXOSC3), Rrp43 (EXOSC8), and Rrp46 (EXOSC5), though only PTM mimetics in Dis3 and Mtr4 impacted RNA processing (119). No studies have explored whether PTMs regulate the function of DDX1. Such studies in the future could provide insight into how the RNA exosome complex dynamically interacts with so many different cofactors to target a large number of distinct RNAs.

Though we have identified and characterized the interaction between the RNA exosome cap subunit EXOSC3 and DDX1, we have not yet explored whether this interaction is direct or indirect. As EXOSC3 is a component of the RNA exosome complex, DDX1 may interact with EXOSC3 or with other RNA exosome subunits. Alternatively, as this interaction was identified through co-purification, the interaction could be indirect and mediated by RNA exosome cofactors or other proteins. While we detect all RNA exosome subunits co-purifying with EXOSC3, these experiments do not demonstrate that DDX1 interacts directly with any RNA exosome subunit beyond EXOSC3. The interaction between EXOSC3 and DDX1 increases significantly upon RNase treatment, particularly RNase A (**Figure 3.3C**), a phenomenon suggested to occur for protein-protein interactions (107). One possible explanation for why the EXOSC3-DDX1 interaction increases upon digestion of RNA may be that RNA is bound between the two proteins, and potentially interfering with or competing for the same binding site. In one conceivable model, DDX1 could unwind RNA for degradation/processing by the RNA exosome and the removal of RNA increases interaction with EXOSC3. A previous study reported a similar observation when examining the interaction between hnRNPK and DDX1 (107). Thus, future studies will be required to further characterize the interaction to test whether DDX1 could function as an RNA exosome cofactor like multiple other helicases.

A logical model to explain the interaction between the RNA exosome and DDX1 was that these factors could cooperate in response to DNA damage. The RNA exosome and DDX1 have both been implicated in double-strand break repair by homologous recombination (HR) and non-homologous end joining (NHEJ) (56,87,100,101). If these factors worked together to respond to DNA damage, we speculated that inducing DNA damage would increase this interaction. In contrast to this prediction, we found that double-strand breaks induced by treatment with the

topoisomerase inhibitor camptothecin significantly reduced the interaction (**Figure 3.4B**). This result suggests that as DDX1 is recruited to sites of DNA damage (101,102), the interaction with the RNA exosome is lost. This finding led us to consider the possibility that the RNA exosome and DDX1 could share a function in cellular homeostasis in the absence of DNA damage. Alternatively, another model that cannot yet be eliminated is that the RNA exosome and DDX1 could cooperate to respond to specific types of DNA damage, which have not yet been tested.

In addition to DNA damage response, both the RNA exosome and DDX1 have been implicated in rRNA processing. The best-defined role for the RNA exosome is 3' trimming to produce mature 5.8S RNA (1,17,83,110), while DDX1 has been implicated in the accumulation of a number of rRNA species (112). Analysis of rRNA processing in cells depleted of either EXOSC3 or DDX1 revealed that loss of these proteins alters some shared rRNA species, albeit in distinct manners (**Figure 3.6**). For example, the mouse precursor of mature 5.8S rRNA, 12S rRNA, accumulates in cells depleted of EXOSC3 with a concomitant decrease in mature 5.8S. In contrast, depletion of DDX1 led to a decrease in the level of the 12S rRNA precursor. While these results show that some of the same rRNAs are impacted by the loss of either EXOSC3 or DDX1, the impact is different.

Helicases belonging to the DEAD- and DExH-box families such as DDX1 and MTR4 play important roles in RNA processing beyond ribosomal RNA. R-loops are also a common target of these helicases. Though R-loops are necessary for cellular maintenance, these structures can pose a threat to the genome if they accumulate (88,120). DEAD/DExH-helicases are critical for resolving and regulating R-loops as they unwind the nucleic acid structures for subsequent degradation by ribonucleases. For example, DDX1 has been reported to unwind G-quadruplex structures that can stabilize R-loops during transcription (100). MTR4, a well-characterized

nuclear cofactor, unwinds R-loops and degrades RNA in complex with the RNA exosome (121). Despite data suggesting that both DDX1 and MTR4 could help resolve R-loops, a number of differences between the helicases exist. For example, the SPRY protein binding domain in the N-terminus upstream of two helicase domains (**Figure 3.3A**) is unique to DDX1. The SPRY domain in DDX1 is inserted between a phosphate-binding P-loop motif and a single-strand DNA binding Ia motif, separating the motifs by 240 residues, instead of the usual 20-40 residues seen in other DEAD-box proteins (96). Examination of the SPRY domain of DDX1 compared with the protein-protein interface of MTR4 and the RNA exosome may reveal surface sites of interaction and would potentially explain why a SPRY domain lies upstream of two helicase domains within the DDX1 protein. Structures indicate that MTR4 interfaces with MPP6, which tethers the helicase to the RNA exosome cap (5,6). This interface could be a shared docking site for these helicases. Further studies on whether DDX1 and MTR4 may have some common functions would also shed light on the cellular roles of these proteins.

Many reports link either the RNA exosome or DDX1 to the modulation of R-loops (87,100,116,117,120). However, no reports to date directly compare the RNA exosome and DDX1 in R-loop modulation. One previous study analyzed a well-characterized R-loop region within the *BAMBI* gene by mass spectrometry and provided a detailed list of the proteins that co-immunoprecipitated with R-loops in this region (116). DDX1, EXOSC3, and several other subunits of the RNA exosome were detected in the association with this *BAMBI* R-loop. Consistent with this previous study, the DRIP-seq performed here identified R-loop regions in the *BAMBI* gene significantly altered in cells depleted of either EXOSC3 or DDX1 (**Figure 3.S3**). These findings support a model where, at least for the *BAMBI* R-loop region, both EXOSC3 and DDX1 are co-located and may contribute to R-loop modulation.

We discovered several genes for which R-loop regions are significantly altered in the same manner by depletion of either EXOSC3 or DDX1. We coupled the genome-wide DRIP-seq analysis with RNA-seq analysis to focus on regions where changes in R-loops are associated with changes in transcript levels. This analysis revealed several common findings. While depletion of either EXOSC3 or DDX1 caused both some increases and some decreases in the number of R-loops detected, data showed larger numbers of decreased R-loop regions as compared to increased R-loop regions (**Figure 3.7A**). Considering these shared changes, GO analysis revealed that the genes located in regions that show increased R-loops map to genes implicated in a variety of cellular responses to stress, which could reflect a requirement for proper RNA exosome and/or DDX1 function to support normal cell physiology. Strikingly, the only GO term enriched more than 5-fold among regions with decreased R-loops is “Anterior/posterior axon guidance” (Fold enrichment > 35). This finding could suggest that these regions of the genome are less accessible and perhaps less actively transcribed when either EXOSC3 or DDX1 is depleted, suggesting a possible link to the neurological disorders that are caused by missense mutations in genes encoding structural subunits of the RNA exosome.

To define potential shared transcript targets of EXOSC3 and DDX1, we also performed RNA-seq analysis using a pipeline that distinguished protein-coding transcripts from other RNAs. This analysis revealed that more transcripts show a significant change (FDR < 0.05) in levels upon depletion of EXOSC3 (3,949 transcripts) as compared to DDX1 (1,702 transcripts). Notably, the percentage of transcripts increased (56% for siEXOSC3; 57% for siDDX1) and decreased (44% for siEXOSC3; 43% for siDDX1) are similar in both cases; in fact, decreased transcripts slightly outnumber increased transcripts for both depletions. Interestingly, the data reveal a large number of shared transcripts that are regulated, suggesting a shared role in modulating a subset of

transcripts or a shared cellular response at the transcript level. To integrate the DRIP-seq and RNA-seq and provide insight into how changes in R-loops could correlate with altered gene expression within the region of the genome where the R-loop is located, we also narrowed the analysis to only consider regions where both R-loops and transcript levels were changed when either EXOSC3 or DDX1 was depleted with a focus on the shared changes. This analysis provided interesting sets of data to consider. One notable point is that the overall patterns of R-loop and transcript changes displayed in the heatmaps (**Figure 3.9B**) appear similar for both EXOSC3 and DDX1 depletions with regions of increased R-loops sharing more increased transcripts and conversely for regions of decreased R-loops. To delve into the data in more detail, we identified altered R-loops linked to RNA metabolism, DNA repair, and/or neurodevelopment. Two examples from this analysis are *Ints6* and *Celf4*. *Ints6* encodes for a protein that is a noncatalytic member of the Integrator complex, a protein complex involved in the processing of snRNAs, resolving lncRNAs, and metabolizing mRNAs (122,123). The Integrator complex, containing INTS6, is involved in transcription termination of lncRNAs that are synthesized at sites of double-strand breaks, also known as damage-induced long noncoding RNAs (dilncRNAs) (122,124). *Celf4* encodes for an RNA binding protein that regulates mRNA stability and binds 3'UTRs (125). *Celf4* has also been linked to neurodevelopmental disorders and expression is enriched in the central nervous system (126,127). These transcripts may be regulated by the RNA exosome and DDX1 in neurons and dysregulation could contribute to the neurological pathology that occurs in many exosomopathy patients. These two examples represent just a subset of the many altered genomic regions and transcripts that could contribute to cellular dysfunction.

The overall goal of this study was to identify RNA exosome interacting proteins present in a neuronal cell line with the underlying hypothesis that neuronal-specific interacting partners could

be lost in exosomopathies, which show primarily neuronal pathologies despite the ubiquitous expression of the RNA exosome. While the identification of an interaction with an additional RNA helicase is exciting, DDX1, like the RNA exosome, is ubiquitously expressed. DDX1 has not been definitively linked to any monogenic disease, but there is a report within the undiagnosed disease network of an individual with a missense mutation (p.Thr280Arg), which would fall within the catalytic domain of DDX1. This individual is reported to suffer from seizures and developmental regression, which could suggest that DDX1, like the RNA exosome, is required for normal function in some regions of the brain. In summary, we identified and characterized a new interacting partner of the RNA exosome, and studies thus far suggest that DDX1 could interact with the RNA exosome to modulate R-loop accumulation at some loci and to regulate transcript levels of some shared transcripts. Taken together, this study suggests the interaction between the RNA exosome and RNA helicase DDX1 may be required for neuronal-specific gene regulation.

Chapter 3 Figures

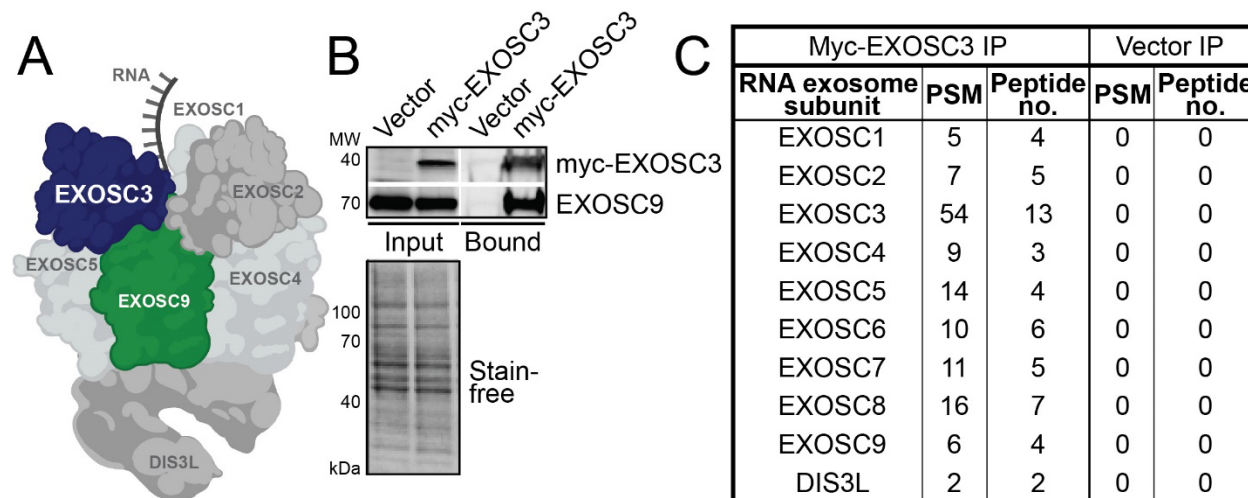


Figure 3.1. RNA exosome subunits co-immunoprecipitate with tagged EXOSC3.

(A) The RNA exosome is a conserved exo/endoribonuclease complex that comprises 10 subunits. Nine of the ten subunits are structural and termed exosome components. EXOSC1, EXOSC2, and EXOSC3 make up the cap and EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, and EXOSC9 make up a barrel-shaped core. In this graphic, EXOSC3 (navy) and EXOSC9 (green) are highlighted. EXOSC6, EXOSC7, and EXOSC8 are positioned behind subunits EXOSC4, EXOSC5, and EXOSC9, and consequently are not visible. The catalytic subunit, DIS3 or DIS3L, sits at the base of the complex. This structure was created using Biorender and is based on PDB 6H25 (4). (B) EXOSC9 core subunit co-precipitates with myc-EXOSC3 from murine neuronal N2A cell line. Cells were transfected with a plasmid encoding Vector control or myc-EXOSC3 followed by immunoprecipitation using anti-myc magnetic beads. Input for Vector control and myc-EXOSC3 was probed by an anti-myc antibody, and a band corresponding to the molecular weight is detected in the Input but not Vector control for myc-EXOSC3. Input for Vector control and myc-EXOSC3 is probed by an anti-EXOSC9 antibody, and a band at the corresponding

molecular weight is present in both lanes. Bound for Vector control and myc-EXOSC3 is probed by an anti-myc and an anti-EXOSC9 antibody, and a band corresponding to the molecular weight is detected in the Bound fraction for myc-EXOSC3 but not for Vector control. Stain-free blot indicates the loading of total protein in the Input. Immunoprecipitation of the myc tagged EXOSC3 copurifies with the endogenous EXOSC9 subunit. (C) Eluates of the Bound myc-EXOSC3 immunoprecipitation were analyzed by LC-MS/MS. A table shows all RNA exosome subunits detected, listing the peptide-spectrum matches (PSM) and the peptide numbers for each subunit. Vector IP serves as a control.

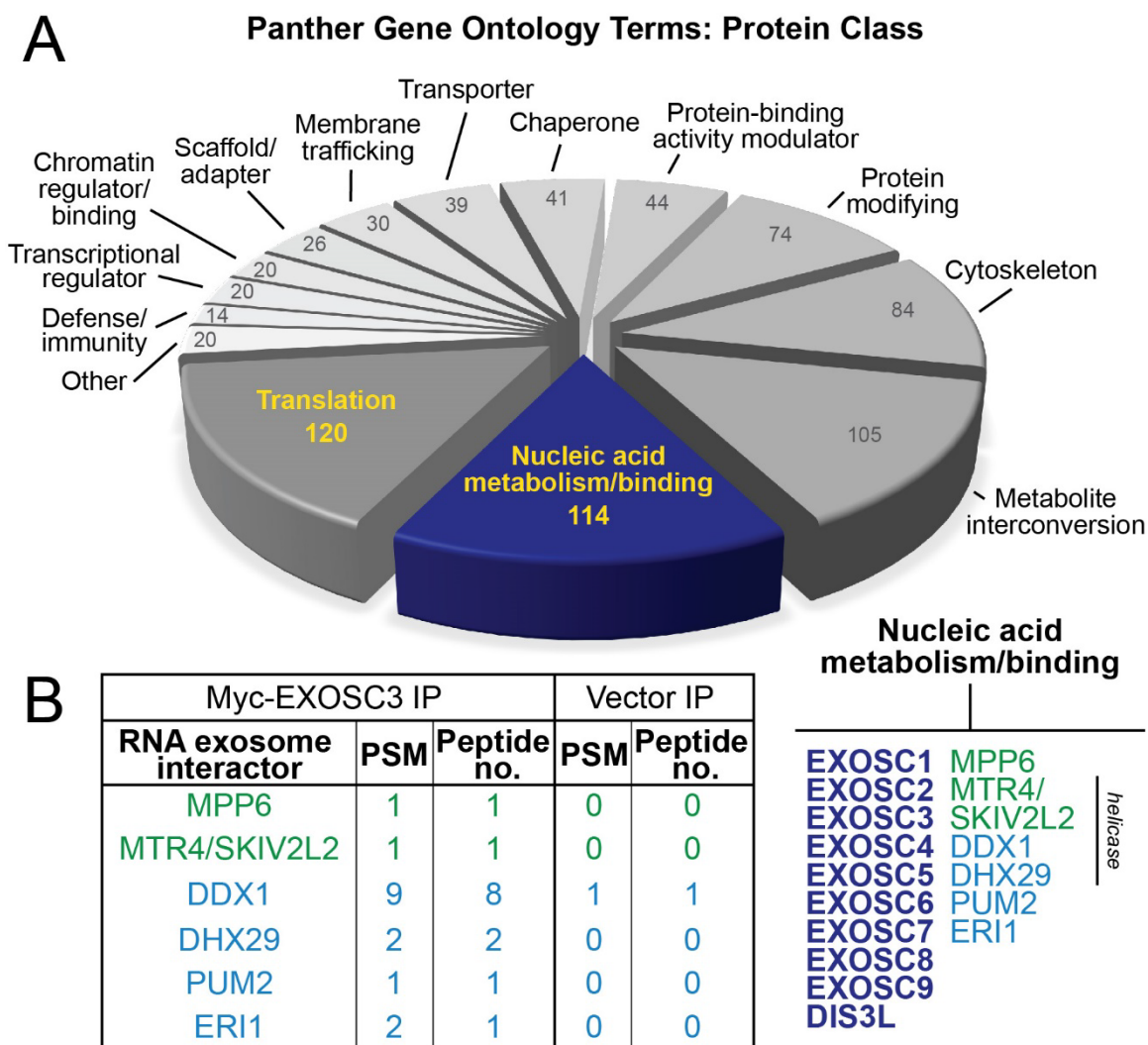


Figure 3.2: Novel EXOSC3/RNA exosome interactors identified using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS).

(A) Pie chart that organizes proteins that co-precipitate with myc-EXSOSC3 from murine neuronal N2A cell line by protein class. Peptide-spectrum matches (PSM) were analyzed using a \log_2 ratio, so that any result above 0 indicates binding in myc-EXOSC3 IP over Vector IP. Results that are less than or equal to 0 were excluded from the analyses. Proteins that contained a PSM beyond the \log_2 cut-off of 0 were analyzed by Panther Gene Ontology terms. The number of proteins within a class is indicated inside the pie slices. (B) The table (left) shows selected RNA exosome interactors

detected and lists the PSM and Peptide number for each protein. Vector IP serves as the control. The gene list corresponding to the proteins is provided as Supporting Information Table S1. Note that MPP6 is the protein name and MPHOSPH6 is the gene name. MTR4 is the most common name for this helicase but appears as SKIV2L2 in Table S1. A short list of Nucleic acid metabolism/binding (right) is provided. The RNA exosome subunits (bold, navy) reside in the nucleic acid metabolism/binding subcategory, together with known RNA exosome cofactors (green). Candidates investigated as novel RNA exosome interactors are listed in blue, including the putative RNA helicase, DDX1.

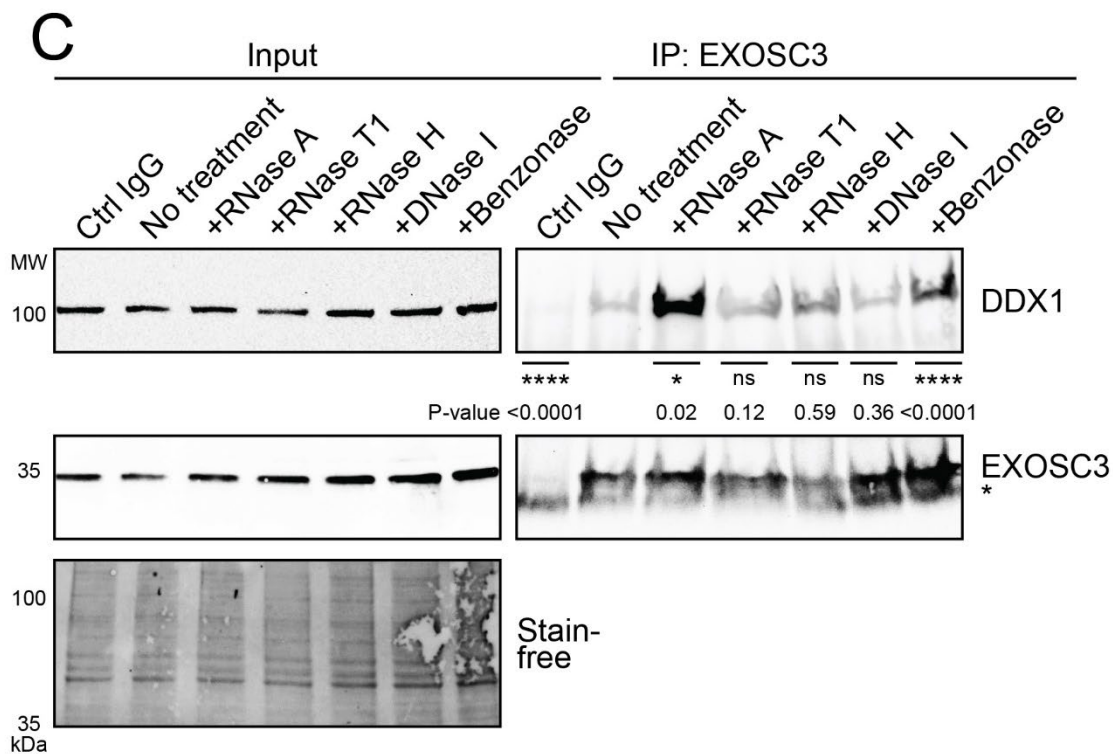
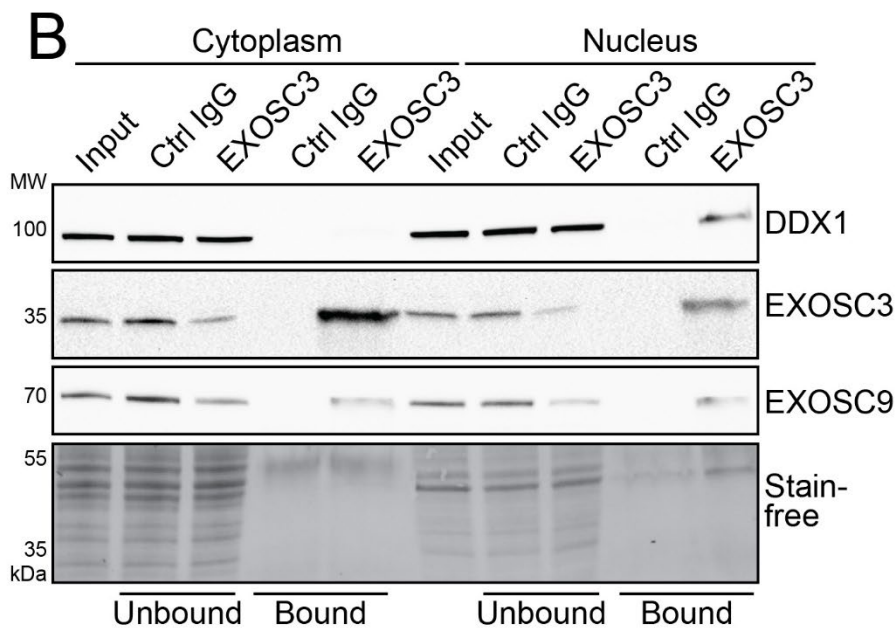
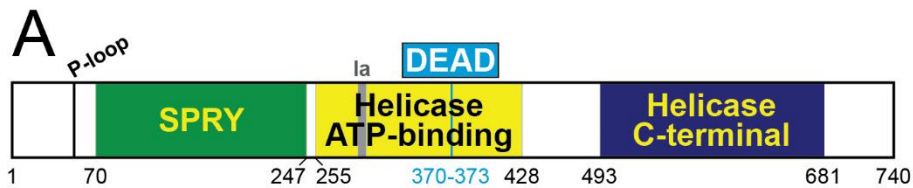


Figure 3.3: DDX1 co-immunoprecipitates with EXOSC3.

(A) A graphical representation of the domain structure of human putative DEAD-box helicase DDX1. The SPRY protein interacting domain in DDX1 is located between a phosphate-binding P-loop motif and a single-strand DNA binding Ia motif, separating the motifs by 240 residues instead of the usual 20-40 residues seen in other DEAD-box proteins (96). The catalytic ATP-binding helicase and C-terminal helicase domains lie downstream of the SPRY domain. (B) DDX1 co-immunoprecipitates with EXOSC3 in the nuclear, but not cytoplasmic, fraction of N2A cell lysate. EXOSC3 was immunoprecipitated from the cytoplasmic or nuclear fraction, followed by immunoblotting using EXOSC3, EXOSC9, and DDX1 antibodies. The Input, Unbound, and Bound fractions from the EXOSC3 and control nonspecific rabbit IgG (Ctrl IgG) immunoprecipitation are shown. EXOSC9 serves as a representative of the co-purified RNA exosome subunits. Stain-free blot serves as the loading control. (C) Immunoprecipitation of EXOSC3 from the nuclear fraction was performed with treatments of RNase A, RNase T1, RNase H, DNase I, and Benzonase. EXOSC3 antibody described previously is used in the No treatment, +RNase A, +RNase T1, +RNase H, +DNase I, and +Benzonase immunoprecipitation. Nonspecific rabbit IgG (Ctrl IgG) was used as a control. DDX1 and EXOSC3 were analyzed for the Input and bound fractions (IP: EXOSC3). The light chain IgG band is visible (asterisk) in the bound fractions probed with EXOSC3, just below the EXOSC3 band. Stain-free serves as a loading control for the Input. Bands in the bound fractions were quantified relative to the No treatment control. The values below the lanes correspond to the amount of protein quantified from the bands. Treatment with RNase A and RNase H were performed in triplicate ($n = 3$). Treatment of RNase T1, DNase I, and Benzonase were performed in duplicate ($n = 2$). The asterisk (*) below the values indicate the P-value < 0.05 .

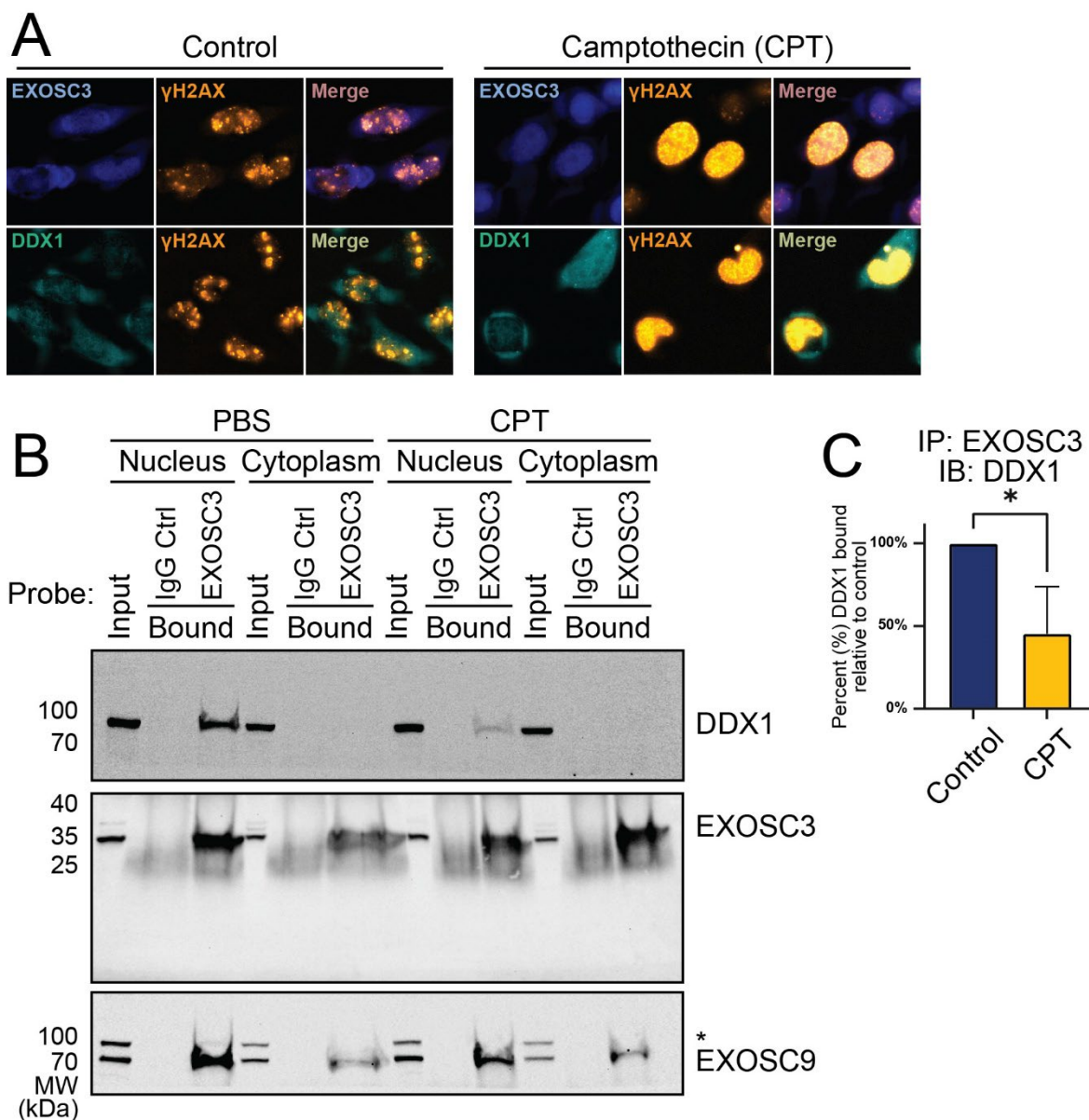


Figure 3.4: The interaction between EXOSC3 and DDX1 is sensitive to DNA damage.

(A) N2A cells were treated with camptothecin (CPT) or PBS (Control), fixed, and analyzed by immunofluorescence using antibodies that detect EXOSC3, DDX1, and the DNA damage marker, γ H2AX. (B) The Input and immunoprecipitated samples from both the cytoplasmic and nuclear fractions (Bound) treated with either CPT or PBS (Control) for both EXOSC3 and control IgG (Ctrl IgG) are shown. DDX1, EXOSC9, and EXOSC3 are detected. (C) The immunoprecipitation

experiment from the nuclear fraction in Figure 4B was performed in biological triplicate and DDX1 bands in EXOSC3 Bound fractions were quantified. Statistical significance was calculated by a student's T-test. Asterisk (*) represents p -value < 0.05 . IP: immunoprecipitation; IB: immunoblot.

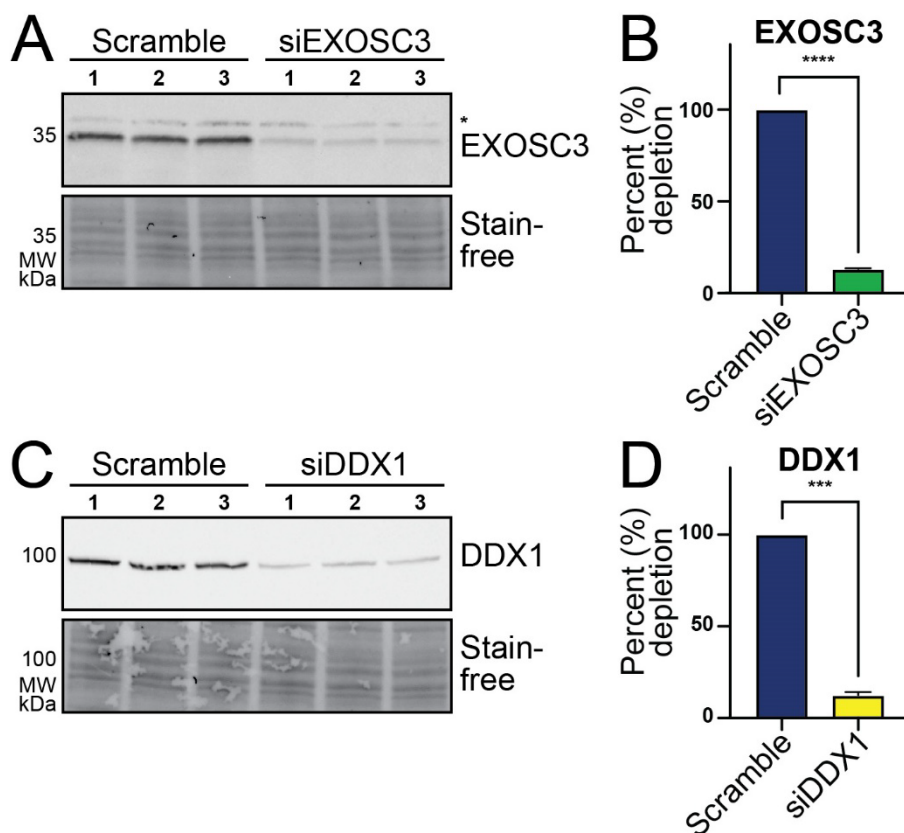


Figure 3.5: EXOSC3 and DDX1 are robustly depleted by siRNA-mediated knockdown in N2A cells.

(A) N2A cells were transfected with Scramble control or EXOSC3 siRNA (siEXOSC3) in biological triplicate. The steady-state level of EXOSC3 was assessed by immunoblotting. The asterisk (*) indicates a nonspecific band. Stain-free blot serves as a loading control. (B) Quantification of immunoblot in (A) shows that EXOSC3 is depleted to 13.1%. The percent depletion of EXOSC3 was determined by quantifying the immunoblot in (A) relative to Scramble and averaging the values. This result is significant across three biological replicates. (C) N2A cells were transfected with Scramble or DDX1 siRNA (siDDX1) in biological triplicate. The steady-state level of DDX1 was assessed by immunoblotting. Stain-free blot serves as a loading control. (D) Quantification of immunoblot in (C) shows that DDX1 is depleted to 12.4%. The percent

depletion of DDX1 was determined by quantifying the immunoblot in (C) relative to Scramble and averaging the values. This result is significant across three biological replicates. The statistical analyses for (C) and (D) were calculated using a student's T-test. Asterisks (****) represent a p-value < 0.0001 and (***) represent a p-value < 0.001 .

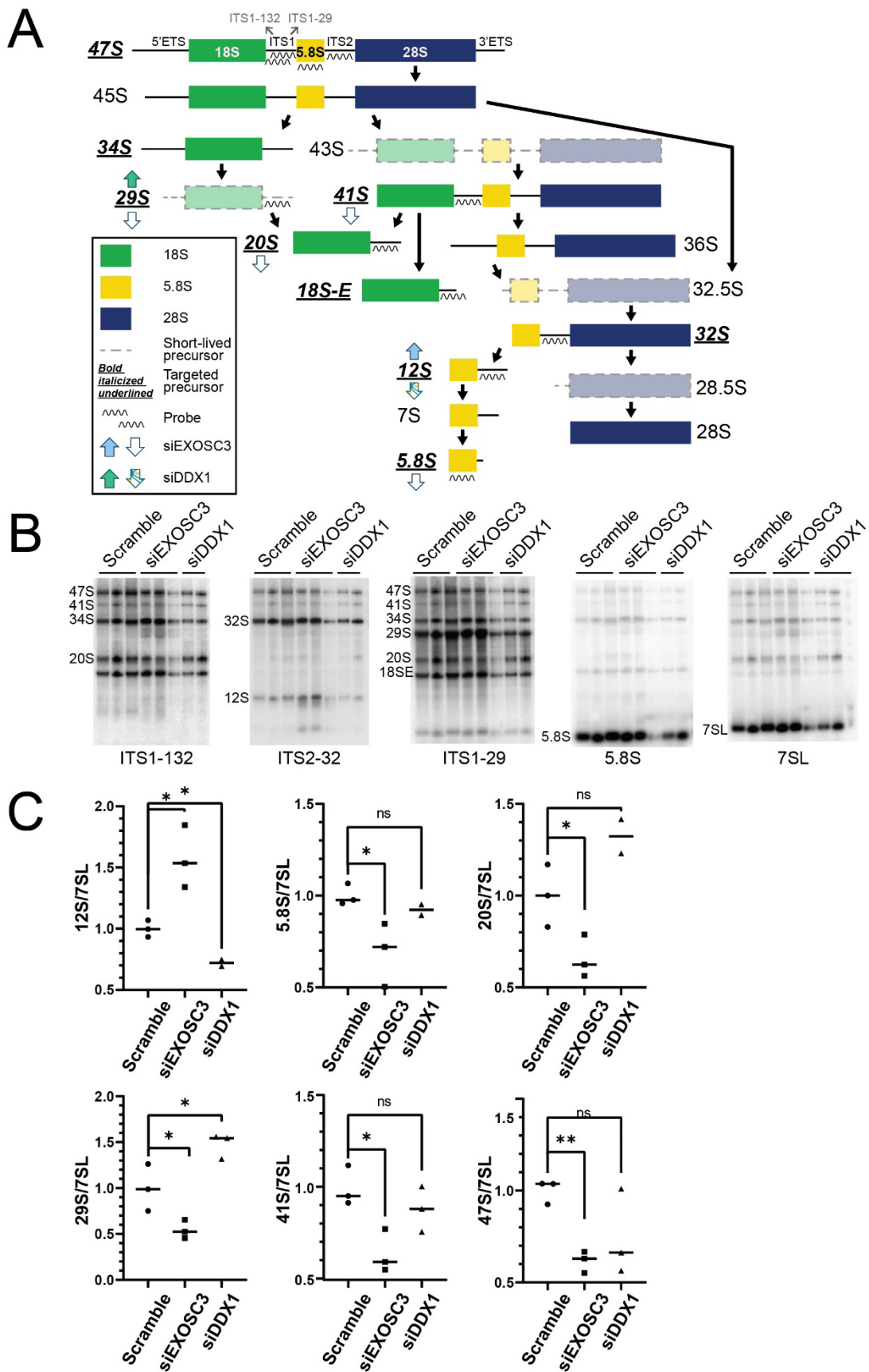


Figure 3.6: Depletion of EXOSC3 or DDX1 results in misprocessing of rRNA precursors.

(A) A graphical schematic of murine rRNA processing, adapted from Henras *et al.*, 2015 (113), that includes up- and down-arrows to summarize the results obtained in this study. (B) Northern blots of total RNA from N2A cells depleted of EXOSC3 or DDX1 using rRNA probes show that levels of 29S, 41S, 20S, and 12S rRNA precursors and 5.8S rRNA are altered. EXOSC3 or DDX1 was depleted from cells by siRNA knockdown and total RNA was isolated for northern blotting. An ITS2 probe was used to detect 32S and 12S rRNA precursors. Additionally, we employed a probe specific for 5.8S rRNA. The 7SL transcript serves as a loading control. (C) Northern blots from Figure 6B were quantified relative to 7SL in biological triplicates. An asterisk (*) indicates a significant difference using a P-value cut-off of < 0.05 .

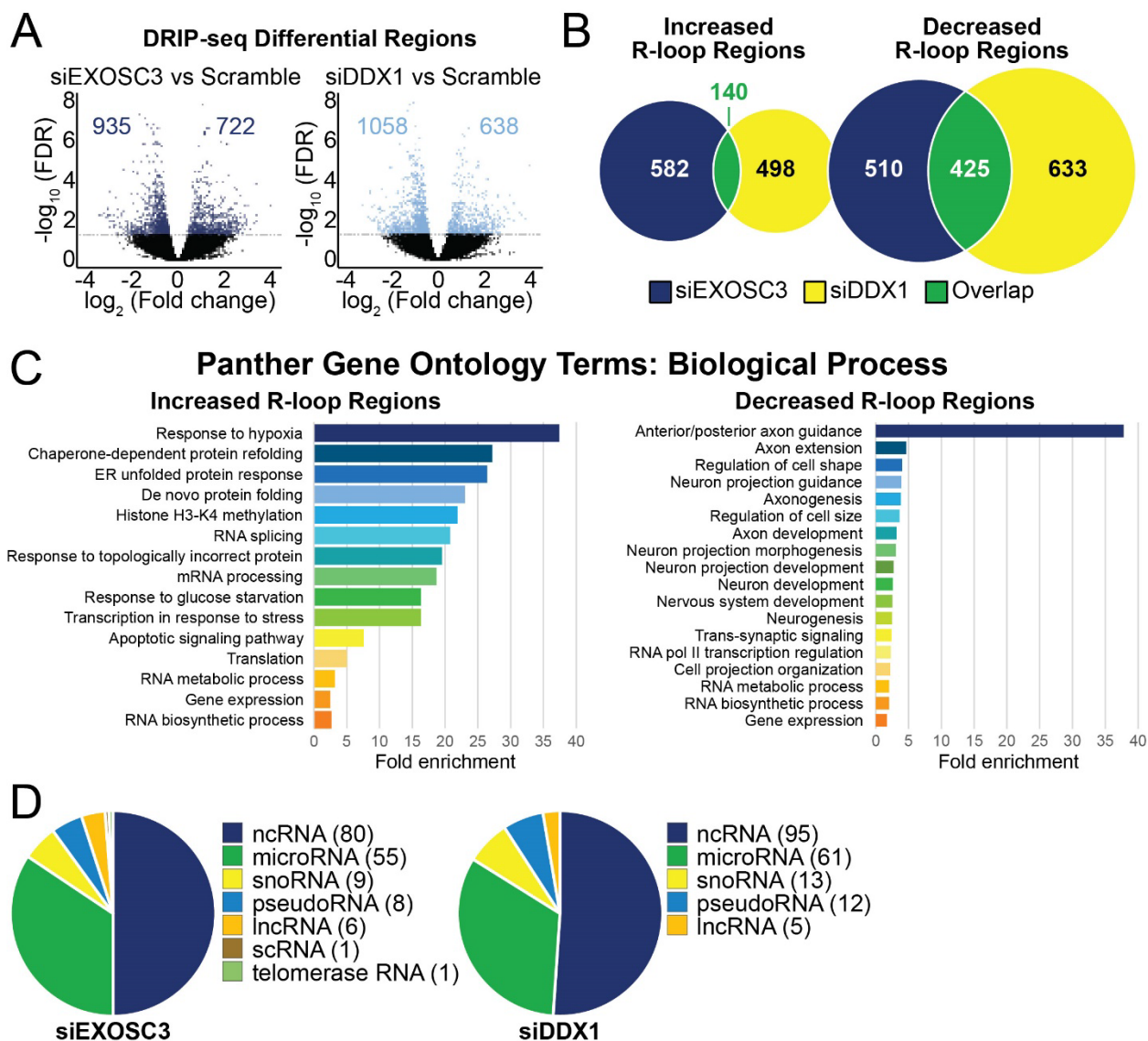


Figure 3.7: DRIP-seq reveals that depletion of EXOSC3 or DDX1 alters R-loop regions.

DNA/RNA immunoprecipitation followed by sequencing (DRIP-seq) was performed on N2A cells depleted of EXOSC3 or DDX1 by siRNA. (A) The volcano plots show the number of R-loop regions that statistically increase and decrease in either EXOSC3 or DDX1 depletions compared to Scramble control. The plot is graphed using a \log_2 fold change across a $-\log_{10}$ false discovery rate (FDR). The R-loop regions that did not achieve the FDR cut-off of 0.05 are indicated in black and fall under the horizontal lines. The numbers on the left of 0 on the x axis indicate significantly

decreased R-loop regions and the numbers on the right of 0 indicate significantly increased R-loop regions. (B) The statistically increased and decreased R-loop regions in cells siRNA depleted of EXOSC3 or DDX1 identified in the DRIP-seq dataset are compared using a Venn diagram. The dark blue circles indicate the number of increased ($n = 722$) or decreased ($n = 935$) R-loop regions upon siRNA-mediated EXOSC3 depletion. The yellow circles represent the number of increased ($n = 638$) or decreased ($n = 1,058$) R-loop regions that were affected upon siRNA-mediated DDX1 depletion. The overlap in green indicates the number of increased ($n = 140$) or decreased ($n = 425$) R-loop regions that siEXOSC3 or siDDX1 have in common. (C) The increased and decreased R-loop regions from both siEXOSC3 and siDDX1 samples were analyzed using Panther Gene Ontology Terms and categorized by biological process. These analyses are statistically significant with a cut-off at $p\text{-value} < 0.05$. (D) The classes of RNA that were affected upon depletion of either EXOSC3 or DDX1 are organized into a pie chart. Protein-coding genes were excluded. The numbers to the right of the class of RNA are the number of R-loop regions that correspond to that class.

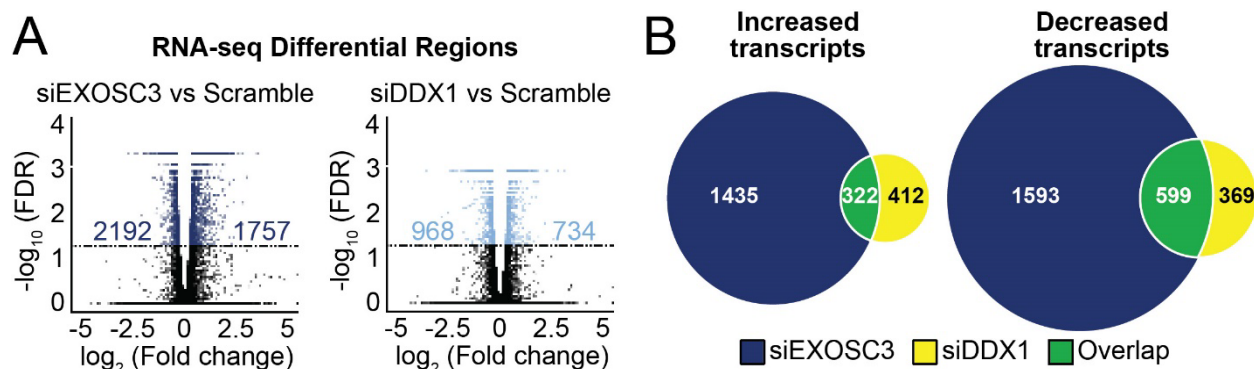


Figure 3.8: RNA-seq shows that depletion of EXOSC3 or DDX1 results in more shared decreased transcripts than shared increased transcripts.

(A) RNA sequencing was performed on N2A cells siRNA depleted of EXOSC3 or DDX1. The volcano plots show the number of mRNA transcripts that increased and decreased in either siEXOSC3 or siDDX1 compared to Scramble control. The plot is graphed using a \log_2 fold change across a $-\log_{10}$ false discover rate (FDR). The differential regions that did not achieve the FDR cut-off of 0.05 are indicated in black and fall under the horizontal lines. The numbers on the left of 0 of the x-axes indicate significantly decreased differential regions and the numbers on the right of 0 indicate significantly increased differential regions. (B) The increased and decreased transcripts are compared using a Venn diagram. The dark blue circles indicate the number of increased ($n = 1,757$) or decreased ($n = 2,192$) mRNA transcripts upon EXOSC3 depletion. The yellow circles represent the number of increased ($n = 734$) or decreased ($n = 968$) mRNA transcripts that are affected upon DDX1 depletion. The overlap represented in green indicates the number of increased ($n = 322$) or decreased ($n = 599$) mRNA transcripts that depletion of EXOSC3 or DDX1 have in common.

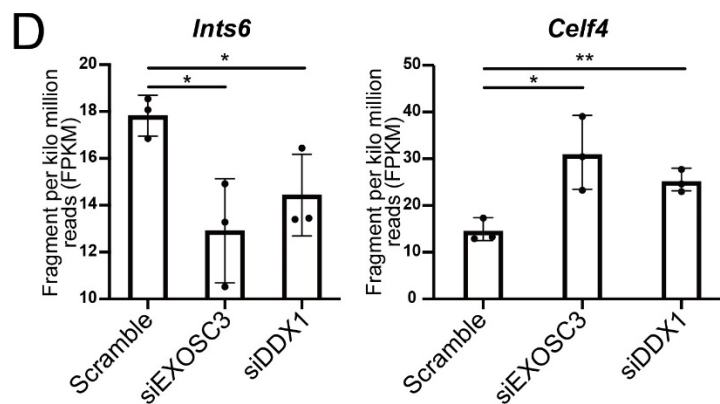
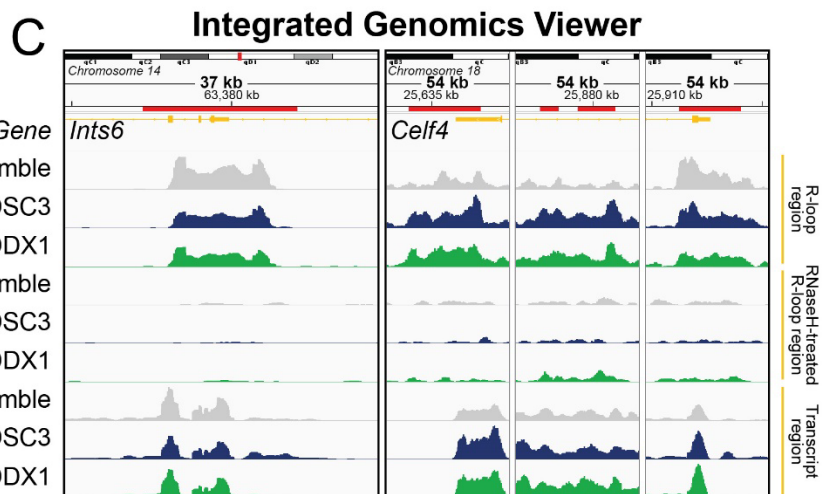
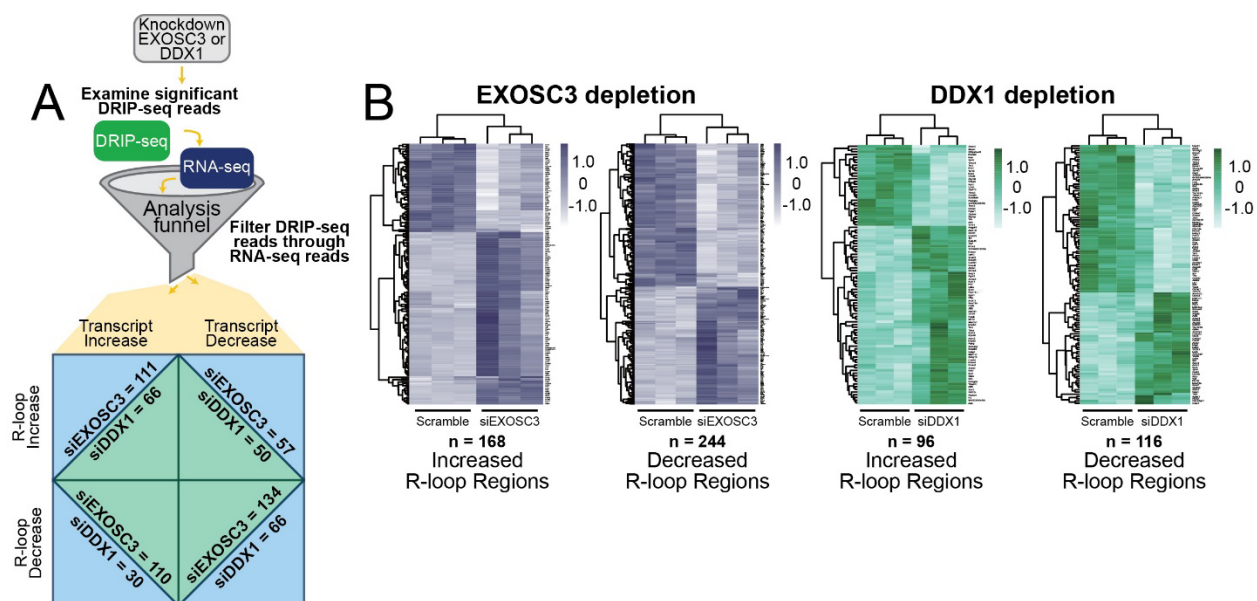


Figure 3.9. Filtering DRIP reads through RNA sequencing revealed genes that are simultaneously affected by depletions of EXOSC3 or DDX1.

(A) A graphical representation of the pipeline used to focus on specific genes using the analysis funnel. N2A cells siRNA depleted of either EXOSC3 or DDX1 were subjected to both DRIP- and RNA-seq. We filtered results from the DRIP-seq through the RNA-seq reads, using only mRNA transcripts. We were then able to produce heatmaps and examine transcriptomic regions using an Integrated Genomics Viewer (IGV). (B) Using the pipeline described, we created heatmaps showing the landscape of increased and decreased R-loop regions upon depletion of either EXOSC3 (blue) or DDX1 (green). (C) The IGV images of *Ints6* and *Celf4*. The chromosome is displayed at the top of the window. The span lists the number of bases currently displayed. The tick marks indicate the chromosome locations. The red line marks the regions in which R-loops are significantly changed. The top track displays the *Mus musculus* reference genome (NCBI37/mm9) in orange. The following three tracks display the R-loop regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively. The middle three tracks display the RNase H-treated R-loop regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively. The last three tracks display the Transcript regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively. *Celf4* exhibited three changed regions in the genes and is displayed by separate panels. (D) Quantification of *Ints6* and *Celf4* transcripts in Scramble, siEXOSC3, and siDDX1 by fragment per kilo million reads (FPKM).

Northern blot probe sequences

5'ETS-end	GGACAGAGAGCGCGAGAGAG
ITS1-29	ACGCCGCCGCTCCTCCACAGTCTCCCGTT
ITS1-132	TTCTCTCACCTCACTCCAGACACCTCGCTCCACA
ITS2-32	ACCCACCGCAGCGGGTGACGCGATTGATCG
28S	CTAATCATTGCTTTACCGG
18S	TAATGATCCTTCCGCAGGTTCCACC
5.8S-113	GCAAGTGCGTTCGAAGTGTC
7SL	CAAACTCCCGTGCTGATCA

Table 3.S1.

The Northern blot shown in Figure 3.6B was produced using the probes listed in this table.

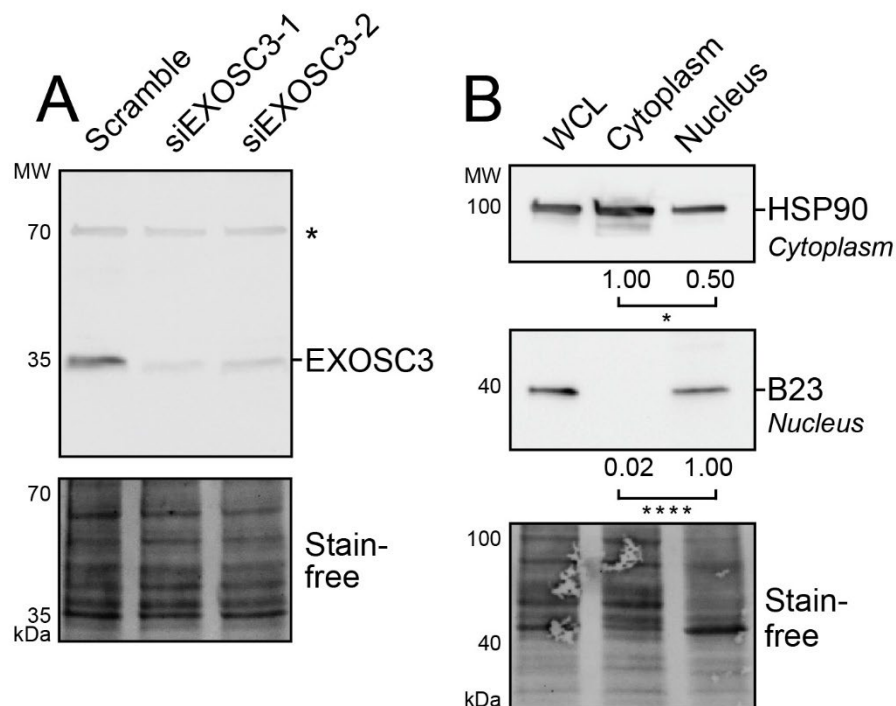


Figure 3.S1. EXOSC3 custom-made antibody is specific, cellular fractionation is sufficient, and the interaction between EXOSC3 and DDX1 is impacted by loss of RNA or DNA.

(A) As described in Materials and Methods, polyclonal antibodies were created to detect EXOSC3. To validate the EXOSC3 antibody, N2A cells were transfected with Scramble or two independent EXOSC3 siRNAs (siEXOSC3-1 and siEXOSC3-2) and protein depletion was determined by immunoblot. The immunoblot was probed by the custom antibody. EXOSC3 antibody detects a band at the predicted size of EXOSC3, which is decreased with each EXOSC3 siRNA compared to Scramble. The Stain-free blot serves as a loading control for total protein. The asterisk (*) denotes a nonspecific band. (B) Cellular fractionation of N2A cells was performed as described in Materials and Methods. Equal amounts of protein from whole cell lysate (WCL), cytoplasmic, or nuclear fractions were analyzed with antibodies that detect proteins localized to either the cytoplasm (HSP90) or the nucleus (B23). The Stain-free blot serves as a loading control. The

differences in protein profiles across the stain-free blot indicate the differences in protein populations between cytoplasmic and nuclear fractions. The bands are quantified relative to the nuclear (B23) or cytoplasmic (HSP90) marker in the respective cellular compartments. The values below the lanes correspond to the amount of protein quantified from the bands, normalized to the fraction observed; the cytoplasmic fraction was normalized to HSP90 (p-value = 0.0279), and the nuclear fraction was normalized to B23 (p-value < 0.0001). The quantification is averaged across biological triplicates and significance was calculated using student's T-test.

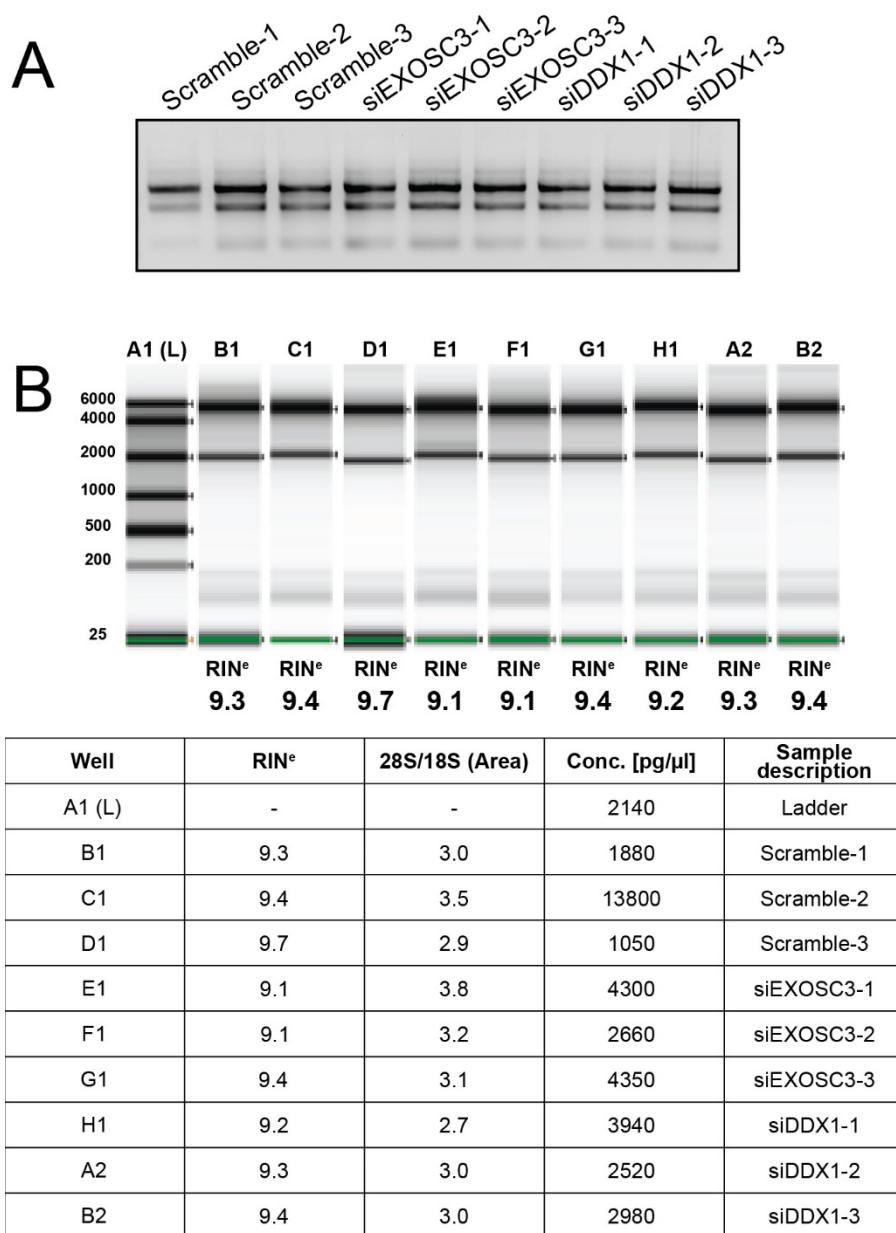


Figure 3.S2. The quality of the purified RNA used for the northern blots, RNA-sequencing, and DRIP-sequencing assessed by 1% agarose gel and High Sensitivity ScreenTape assay.

(A) RNA was extracted from N2A cells transfected with scramble siRNA, EXOSC3 siRNA, and DDX1 siRNA in biological triplicates. The quality of the RNA was assessed by a 1% agarose gel. The three bands visible correspond top to bottom to 28S, 18S, and 5.8S rRNA, respectively. (B)

A High Sensitivity RNA ScreenTape assay was used for quantitation and further quality assessment. All RNA integrity numbers (RIN^o) for each sample are considered high quality.

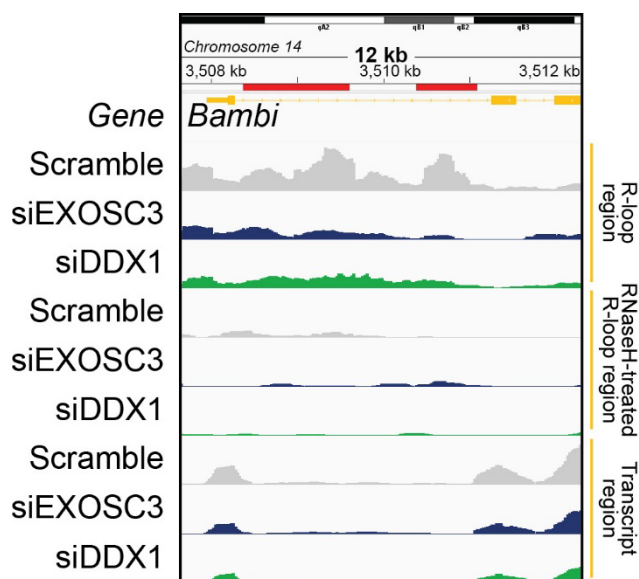


Figure 3.S3. IGV for R-loop regions in *Bambi* is reduced in cells depleted of EXOSC3 or DDX1.

The Integrative Genome Viewer (IGV) image of *Bambi* is shown. The chromosome is displayed at the top of the window. The span lists the number of bases currently displayed. The tick marks indicate the chromosome locations. The red line marks the regions in which R-loops are significantly changed. The top track displays the *Mus musculus* reference genome (NCBI37/mm9) in orange. The following three tracks display the R-loop regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively. The middle three tracks display the RNase H-treated R-loop regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively. The last three tracks display the Transcript regions of interest corresponding to Scramble, siEXOSC3, and siDDX1, respectively.

Conclusions and Future Directions

The studies described in this dissertation explore possible biological basis for a class of diseases termed “exosomopathies”. These works provide key insights to the function of the RNA exosome adds to our model suggesting how pathogenic single amino acid changes within structural subunits of the RNA exosome contribute to neurological disease (**Figure 4.1**). The analysis here supports the characterization of the RNA exosome as ubiquitously expressed across tissues. We find evidence that a key subunit of the complex, EXOSC3, could be rate limiting for assembly in multiple tissues. Furthermore, we identify a potential novel cofactor, DDX1, supporting a model where many cofactors that are yet to be defined could be critical for RNA exosome function.

The RNA exosome has been studied for decades, as the complex was initially defined genetically in budding yeast. Although important discoveries have been made using a variety of genetic model organisms, the recently discovered link to neurological disease raises important questions. For example, yeast are unicellular and cannot provide details of how a genetic mutation may affect different tissues. Although *Drosophila*, are multi-cellular with a complex brain, flies lack brain structures affected in exosomopathies that exist in humans, such as the cerebellum. The *Drosophila* RNA exosome also lacks EXOSC8, a core structural component that has been linked to PCH1C (9). A logical next step beyond model organisms is to use mammalian cell lines. The work described in Chapter 3 used N2A neuronal mouse cell lines to investigate novel interactions with the RNA exosome. Future studies may employ human neuronal cell lines, such as SH-SY5Y, or differentiated neuronal cells to further mimic the cells impacted by mutations causing exosomopathies. Rodent brain, specifically the cerebellum and/or pons, may also be exploited to investigate novel interactions. An approach to further explore the consequences of single amino acid changes in structural RNA exosome subunits would be to employ CRISPR to edit the cell lines described. DDX1 is present in the mammalian brain and is detected in SH-SY5Y cells (128),

and therefore further molecular nuances of the interaction between the RNA exosome and the novel helicase interactor may be explored in these systems.

Exosomopathies are very rare diseases, and therefore patient samples are quite limited. A mouse model would be a beneficial system to study molecular consequences of mutations in structural RNA exosome subunit genes. To date, no mouse model has been created for any exosomopathy variant. One study used an *ex vivo* approach by employing a conditional inversion method ultimately resulting in an EXOSC3 knockout from cultured B cells in adult mice (129). The Jackson Laboratory lists CRISPR-generated knockout lines for *EXOSC1* and *EXOSC2* for mice, but there are no reports to date that indicate use of these lines or analysis of these strains. Moreover, specific edits to the genome of a mammalian model may provide certain insights that would be lost in a knockout system. Now, with a breadth of data from studies with other genetic models, the choice of variants to proceed with in producing that mouse model is well informed. The goals of these models are to understand the biological consequences that underlie exosomopathies and ultimately to test novel therapeutics.

One explanation for why a mouse model has not yet been created to date is the likelihood of survival. Knockout mice would likely not be viable. Even patients, who predominantly have missense mutations, often do not survive puberty and many die in infancy. Another method of testing biological consequences in a multi-cellular mammalian system is by the creation of organoids. Organoids are three-dimensional tissues typically derived from stem cells that mimics organs (130). An organoid system that mimics the cerebellum containing genome-edited exosomopathy mutant variants would be an invaluable model for the RNA exosome field. An *in vitro* organoid system containing cerebellar cells, including Purkinje and granule cells (53), would facilitate testing cell growth and development, cell morphology, changes in RNA and/or protein

expression, changes in the transcriptome, and impacts on R-loops and/or sites of transcription. Additionally, cell-specificity and/or cell specific protein-protein interactions between the RNA exosome and potential cofactors has not yet been determined and novel interactors are continuously discovered (66). These approaches may provide insight to the specific cell types most impacted by exosomopathies.

Exosomopathies and the biological consequences of pathogenic missense mutations in RNA exosome subunit genes.

Nearly all patients with exosomopathies present with cerebellar hypoplasia and/or progressive cerebellar atrophy with varying severity, and many of them are diagnosed with a specific subtype of PCH. Clinical reports of exosomopathy patients usually rely on an MRI of the brain and genetic testing. Other tissues seem to develop normally. While the pons and cerebellum are clearly impacted, clinical pathologies report functional abnormalities in tissues beyond the pons and cerebellum, including the brain cortex, spinal cord, heart, kidney, lung, muscle, skin, and thyroid. The most notable abnormalities could be the result of underdevelopment and/or degeneration of the pons and cerebellum. The pons and pontine circuits are crucial for controlling respiration (131), which may explain why so many patients require breathing support. The medulla, where the brain meets the spinal cord, is in close proximity to the pons, and plays major roles in blood circulation (132), which could be why patients have experienced hypertension. The cerebellar nuclei project to the ventrolateral thalamus, which projects to many cortical areas, including areas of the frontal, prefrontal, and posterior parietal cortex (133). Novel neural circuits are routinely discovered, and therefore, we have yet to understand the full extent of the effects of

the loss of cerebellar neurons and the biological consequences of pathogenic missense mutations that cause exosomopathies.

In Chapter 2, where we analyze gene expression of RNA exosome subunits, data show high RNA exosome transcript levels across the cerebellar hemisphere and cerebellum as compared to other tissues examined. These results suggest that these tissues could require more RNA exosome for cellular homeostasis compared with the heart, muscle, and kidney. These results could begin to explain why patients with missense mutations in structural RNA exosome subunit genes present with cerebellar mass loss and have the most well-defined phenotypes in the cerebellum. Although the transcript levels do not necessarily correlate with the amount of protein in cells or tissues (134), we suspect these levels are biologically relevant and required to support proper cell function.

Furthermore, the analysis in Chapter 2 reveals that the *EXOSC3* transcript is lowest of all the RNA exosome subunits across tissues examined, except for kidney. Even for the kidney, the levels of *EXOSC3* are comparatively low in transcripts per million (TPM) (lowest TPM = 5.29, highest TPM = 19.42, *EXOSC3* TPM = 5.53) (**Figure 2.2A**). This observation could suggest that *EXOSC3* is limiting for RNA exosome assembly. These limits could explain why the *EXOSC3* gene was the first RNA exosome subunit gene identified and linked to human disease (79), as even minor changes that alter the function or level of *EXOSC3* could be pathogenic. To date, the largest number of patients with exosomopathies reported have missense mutations in *EXOSC3* (76). Alternatively, this larger number of cases may represent ascertainment bias because *EXOSC3* was the first RNA exosome gene linked to disease; thus, the *EXOSC3* mutations have had the longest time to accrue. Time will tell whether the rare cases of *EXOSC* gene mutations will become more common now that the link to pathology has been established.

As more patients with exosomopathies are identified, types of pathologies associated with mutations in *EXOSC* genes will likely be elucidated. In addition, further proteomic analysis of RNA exosome subunits in specific tissues and cell types, such as Purkinje and granule cells known to be abundant in the cerebellum, could provide insight into why mutations in *EXOSC* genes cause cerebellar pathology as one of the most prevalent phenotypes.

Novel interactions with the RNA exosome in neuronal cells.

One hypothesis for why single amino acid changes could have tissue-specific consequences could be the cell-specific interaction with cofactors are impacted. In Chapter 3, we identified and characterized a novel interaction between EXOSC3 and a putative RNA helicase, DDX1, in a neuronal cell line. We discovered this interaction specifically occurs in the nucleus, despite the fact that the RNA exosome subunit EXOSC3 and DDX1 are present at approximately equal levels in both the nucleus and cytoplasm. Based on our analysis of shared functions, our results suggest that EXOSC3 and DDX1 participate in regulation of R-loops with consequences for transcripts produced from the genomic regions that form those R-loops. Taken together, these data define a potential mechanism by which an interaction between the RNA helicase and the RNA exosome could modulate R-loops.

Further questions remain to be answered regarding the interaction between EXOSC3 and DDX1. We next would explore how the interaction is impacted by the pathogenic mutations present in exosomopathies. Using methods similar to those in **Figure 3.1B**, we could transfect N2A cells with a plasmid encoding a myc-*EXOSC3* mutation (W237R, G31A, or D131A), with simultaneous knockdown of endogenous EXOSC3 by targeting an untranslated region (UTR). Such experiments could help to define the functional consequences of pathogenic mutations.

We have not yet determined whether the interaction between EXOSC3 and DDX1 is direct or indirect. DDX1 may interact directly with EXOSC3 or DDX1 could interact with other subunits of the RNA exosome complex. An assembled RNA exosome complex could be required for the interaction with DDX1. We could test this idea by performing sucrose gradient centrifugation and immunoblotting to determine if DDX1 interacts with the entire complex, a few subunits, or one specific subunit. We also have yet to determine whether other proteins are required for the interaction between EXOSC3 and DDX1. We detect a significant increase in the interaction between EXOSC3 and DDX1 upon treatment with RNase A (**Figure 3.3C**), a result which has been used to suggest a direct interaction (107). DDX1 and RNA could compete for the same binding site on EXOSC3. A similar observation was made for hnRNPK and DDX1 (107). One possible model to explain the interaction between EXOSC3 and DDX1 is that the helicase unwinds some specific target RNA species to thread through the barrel of the RNA exosome for degradation. This model has been demonstrated for other helicases, including MTR4, by elegant structural and biochemical studies (135-137).

Helicases belonging to the DEAD- and DExH-box families such as DDX1 and MTR4 play important roles in RNA processing. R-loops are also a common target of these helicases. R-loops are necessary for cellular maintenance, such as forming transcription bubbles; however, these structures can pose a threat to the genome if they accumulate (88,120). DEAD/DExH-helicases are critical for resolving and regulating R-loops as they unwind the nucleic acid structures for subsequent degradation by ribonucleases. For example, DDX1 has been reported to unwind G-quadruplex structures that can stabilize R-loops during transcription (100). MTR4, a well-characterized nuclear RNA exosome cofactor, unwinds R-loops and degrades RNA in complex with the RNA exosome (121). Despite data suggesting that both DDX1 and MTR4 could help

resolve R-loops, a number of differences between the helicases exist. For example, the SPRY protein binding domain in the N-terminus upstream of two helicase domains (**Figure 3.2A**) is unique to DDX1. The SPRY domain in DDX1 is inserted between a phosphate-binding P-loop motif and a single-strand DNA binding Ia motif, separating the motifs by 240 residues, instead of the usual 20-40 residues seen in other DEAD-box proteins (96). Examination of the SPRY domain of DDX1 compared with the protein-protein interface of MTR4 and the RNA exosome may reveal surface sites of interaction and would potentially explain why a SPRY domain lies upstream of two helicase domains within the DDX1 protein.

Structural studies reveal that MTR4 interfaces with MPP6, which tethers the helicase to the RNA exosome cap (6,137). This interface could be a shared docking site for these helicases on the RNA exosome. Further studies to explore whether DDX1 and MTR4 may have some common functions would also shed light on the cellular roles of these proteins. Other helicases may also interact with the RNA exosome. Several helicases were detected in the studies described in Chapter 3 (**Figure 3.2**). These helicases are present in N2A cells; however, tissue- or cell-specific helicases may interact with the RNA exosome in other contexts, such as in Purkinje cells present in the cerebellum or other tissues such as the heart or the lung. The interaction between the RNA exosome and alternate helicases may be required to process and/or degrade specific RNAs, potentially due to the specific sequences within the RNAs, the composition of the nucleic acids (RNA-DNA or RNA-RNA), the size of the RNA, the structure of the RNA such as the G-quadruplex described previously (100), or the localization of the RNA within the cell.

Helicases and other well-defined RNA exosome cofactors are typically specific to either the nucleus or the cytoplasm. In Chapter 3, we show that DDX1 is present in both the nucleus and the cytoplasm, but the interaction with EXOSC3 is detected only within nuclear fractions. One

explanation for why the interaction between EXOSC3 and DDX1 is compartment-specific is that post-translational modifications (PTMs) could modulate binding. Studies in *S. pombe* have revealed PTMs in several RNA exosome subunits (119); however, no studies to date have explored whether PTMs regulate the function of the RNA exosome or DDX1 in mammals. Studies such as these could provide insight into how the RNA exosome complex dynamically interacts with so many different cofactors to target many distinct RNAs.

One avenue we explored in the work is the potentially cooperative roles the RNA exosome and DDX1 play in ribosome maturation via generation of mature and fully processed rRNA by targeting specific rRNA precursors in mouse rRNA processing in a neuronal cell line. Depletion of either EXOSC3 or DDX1 affects similar precursors, albeit in opposite ways. For example, while depletion of EXOSC3 in neuronal cells increases the 12S precursor, depletion of DDX1 decreases the same precursor (**Figure 3.6**). One possible explanation for why EXOSC3 and DDX1 have inverse effects on the same precursors is that DDX1 could interact with factors, such as other ribonucleases, in the rRNA maturation pathway that participate in producing a number of precursors that promote cell homeostasis. Additionally, the knockdown of either EXOSC3 or DDX1 has an impact on hundreds or even thousands of transcripts (**Figure 3.7** and **Figure 3.8**). Those changes in transcripts may have downstream consequences that affect the maturation of rRNA. We understand the primary role of the RNA exosome in rRNA maturation. To determine why loss of EXOSC3 or DDX1 differentially affected levels of rRNA precursors, the role that DDX1 plays in rRNA maturation must be better defined, such as investigating whether DDX1 interacts directly with ribosomal precursors and which ones.

Although EXOSC3 and DDX1 individually impact rRNA production, ribosomes are likely to be functional in exosomopathy patients, as mature ribosomes are critical for cell survival. *DDX1*

is not linked to a specific monogenic disease, but has been reported to be naturally overexpressed in several neuroblastoma tumors and cell lines (106,138). Indeed, we speculated that the RNA exosome and DDX1 could interact to coordinate DNA repair. Both the RNA exosome and DDX1 have been implicated in double-strand break repair by homologous recombination (HR) and non-homologous end joining (NHEJ) (87,100,101,129). If these factors indeed cooperated in response to DNA damage, we speculated that inducing DNA damage would increase this interaction. In contrast to this prediction, we found that double-strand breaks induced by treatment with the topoisomerase inhibitor camptothecin significantly reduced the interaction (**Figure 3.4B**). This result suggests that as DDX1 is recruited to sites of DNA damage (101,102), the interaction with the RNA exosome is lost. This finding led us to consider the possibility that the RNA exosome and DDX1 could share a function in cellular homeostasis in the absence of DNA damage. Alternatively, another model that cannot yet be eliminated is that the RNA exosome and DDX1 could cooperate to respond to specific types of DNA damage, which have not yet been tested.

We have explored the interaction between EXOSC3 and DDX1 in several ways. One major question left to answer is whether we can consider DDX1 as an RNA exosome cofactor. To qualify as a cofactor, the factor must influence processing and/or degradation activity of the RNA exosome. To address this point, multiple studies would need to be performed. Cofactors have been confirmed by *in vitro* binding assays, RNA processing/degradation activity assays, and a structure solved in complex with the RNA exosome. To determine whether DDX1 is a cofactor, all these experiments must be performed, and the experiments may require significant optimization, such as the presence of other cofactors or specific species of RNA. We speculate that this will be case because despite the fact that both EXOSC3 and DDX1 are present in the nucleus and cytoplasm,

the interaction occurs in the nucleus. Therefore, there may be a nuclear factor required for this interaction.

In this work, we described the current model systems used to explore RNA exosome biology, provided RNA-sequencing analysis for the broad claim that the RNA exosome is ubiquitously expressed, and identified and characterized a novel interaction between EXOSC3 and DDX1 in a neuronal cell line. In our model (**Figure 1.2A**), mutations that give rise to single amino acid substitutions in structural RNA exosome subunits may alter the integrity of the RNA exosome complex by interfering with assembly or making the assembled complex less accessible (**Figure 1.2B**), decrease interactions with specific cofactors (**Figure 1.2C**), or disrupt RNA processing/degradation (**Figure 1.2D**), and. We extended this model by identifying a new RNA exosome interacting protein, DDX1 (**Figure 4.1**). The findings support the notion that the RNA exosome is ubiquitously expressed and a cap subunit of the RNA exosome, EXOSC3, may be limiting for assembly in multiple tissues (**Figure 4.1B**). Furthermore, the amount of RNA exosome subunit transcripts are high in the cerebellum, an area notably impacted in PCH patients, compared to other tissues, possibly suggesting a high requirement of the RNA exosome complex in this brain region. Therefore, a single amino acid change in EXOSC3 could logically affect the levels of the RNA exosome complex in the cerebellum.

This work additionally identifies and characterizes a novel interaction with the RNA exosome and a putative RNA helicase, DDX1 in neuronal cells (**Figure 4.1C**). Like the RNA exosome, DDX1 is a ubiquitously expressed. We have yet to determine if the interaction between the RNA exosome and DDX1 is specific to neuronal cells or if the RNA exosome and DDX1 cooperate to process/resolve specific neuronal transcripts. The GTEx Portal described in Materials and Methods in Chapter 2 reports high levels of DDX1 in the cerebellum compared to other tissues.

The TPM for DDX1 transcripts in the cerebellum is five- to ten-fold the TPM for the RNA exosome subunit transcripts. The fact the DDX1 transcripts are high in the cerebellum could indicate that the interaction between the RNA exosome and DDX1 is important in this tissue. The identification of the potential novel cofactor, DDX1, suggests that other yet-to-be-defined cofactors could be critical for RNA exosome function.

Finally, this work investigated shared R-loops upon depletion of either EXOSC3 or DDX1 (**Figure 4.1D**). We discovered several genes for which R-loop regions are significantly altered. GO analysis of these shared changes revealed that depletion of either EXOSC3 or DDX1 impact biological processes related to neuronal development, which could suggest the interaction between the RNA exosome and the putative RNA helicase DDX1 may be necessary for neuronal-specific gene regulation.

The studies described in this dissertation provide valuable insights into the biological basis of exosomopathies and sheds light on the function of the RNA exosome and its potential role in neurological diseases. These findings pave the way for future research aimed at unraveling the intricate mechanisms underlying exosomopathies, and hopefully developing novel therapeutics to improve the quality of life for patients with these debilitating disorders.

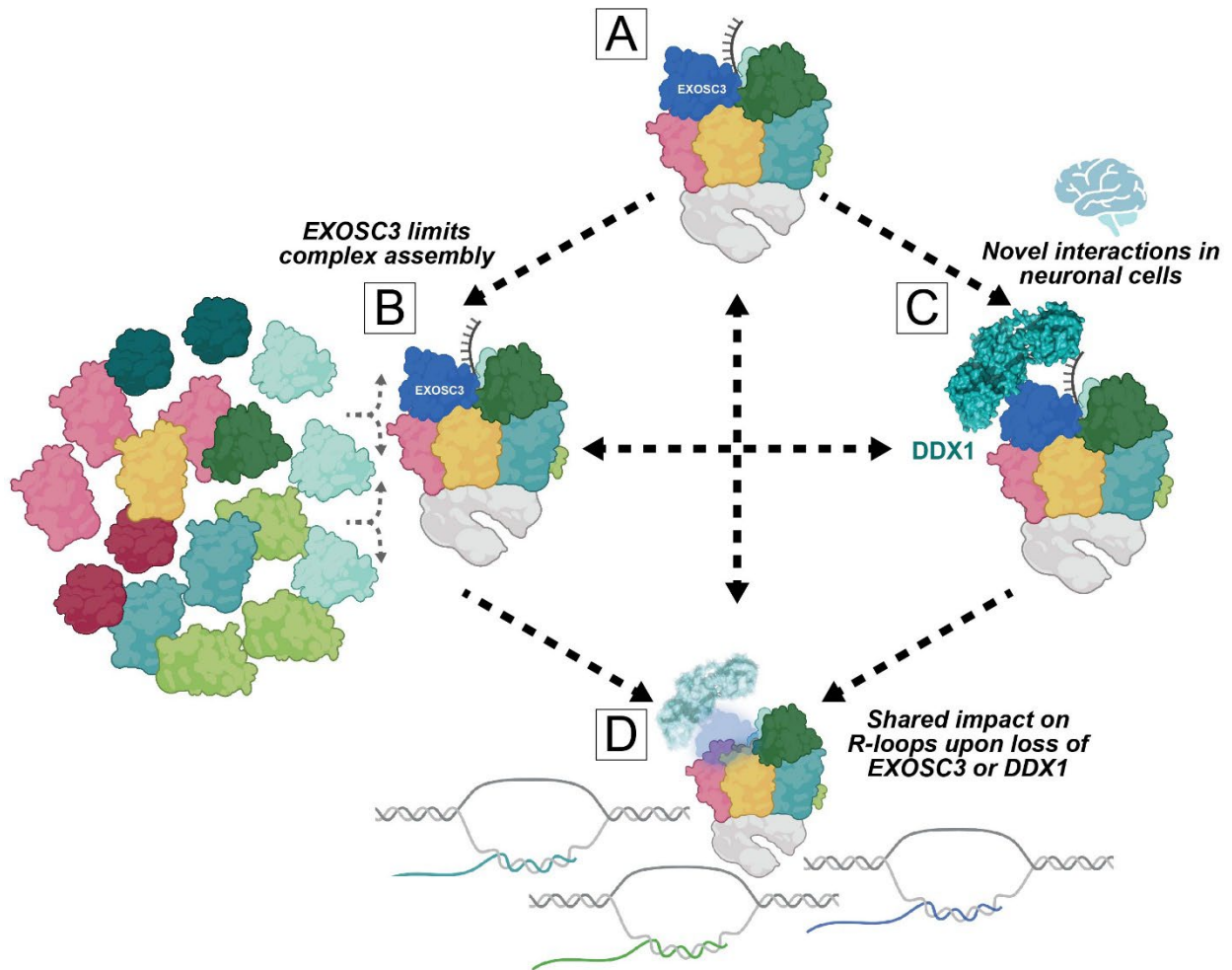


Figure 4.1. Model of how the RNA exosome may impact gene expression regulation and neurological pathology.

(A) The RNA exosome is conserved, ubiquitously expressed across tissues and cell types, and critical for cellular function. (B) Chapter 2 provides evidence that the RNA exosome subunit transcripts are ubiquitously expressed across tissues and that EXOSC3 may be a limiting factor for complex assembly. (C) Data described in Chapter 3 work provide evidence that the novel interactor, DDX1, may cooperate with the RNA exosome to maintain cellular homeostasis (66). Additionally, new potential cofactors may interact with the RNA exosome in a cell- or tissue-specific manner. (D) The work described in Chapter 3 analyzed many shared R-loops upon loss of either EXOSC3 or DDX1. Many of the shared R-loops impacted are involved in processes that aid

in axon guidance, suggesting a possible link to the neurological disorders caused by missense mutations in EXOSC genes.

References

1. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell* **91**, 457-466
2. Zinder, J. C., and Lima, C. D. (2017) Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Gene & Development* **31**
3. Makino, D. L., Baumgärtner, M., and Conti, E. (2013) Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* **495**, 70
4. Gerlach, P., Schuller, J. M., Bonneau, F., Basquin, J., Reichelt, P., Falk, S., and Conti, E. (2018) Distinct and evolutionary conserved structural features of the human nuclear exosome complex. *eLife* **7**
5. Weick, E. M., Puno, M. R., Januszyk, K., Zinder, J. C., DiMattia, M. A., and Lima, C. D. (2018) Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex. *Cell* **173**, 1663-+
6. Falk, S., Bonneau, F., Ebert, J., Kogel, A., and Conti, E. (2017) Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase. *Cell Rep* **20**, 2279-2286
7. Zinder, J. C., Wasmuth, E. V., and Lima, C. D. (2016) Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3. *Mol Cell* **64**, 734-745
8. Francois-Moutal, L., Jahanbakhsh, S., Nelson, A. D. L., Ray, D., Scott, D. D., Hennefarth, M. R., Moutal, A., Perez-Miller, S., Ambrose, A. J., Al-Shamari, A., Coursodon, P., Meechoovet, B., Reiman, R., Lyons, E., Beilstein, M., Chapman, E., Morris, Q. D., Van Keuren-Jensen, K., Hughes, T. R., Khanna, R., Koehler, C., Jen, J., Gokhale, V., and Khanna, M. (2018) A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B). *ACS chemical biology* **13**, 3000-3010
9. Boczonadi, V., Muller, J. S., Pyle, A., Munkley, J., Dor, T., Quartararo, J., Ferrero, I., Karcagi, V., Giunta, M., Polvikoski, T., Birchall, D., Princzinger, A., Cinnamon, Y., Lutzkendorf, S., Piko, H., Reza, M., Florez, L., Santibanez-Koref, M., Griffin, H., Schuelke, M., Elpeleg, O., Kalaydjieva, L., Lochmuller, H., Elliott, D. J., Chinnery, P. F., Edvardson, S., and Horvath, R. (2014) EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. *Nature communications* **5**, 4287
10. Burns, D. T., Donkervoort, S., Muller, J. S., Knierim, E., Bharucha-Goebel, D., Faqeih, E. A., Bell, S. K., AlFaifi, A. Y., Monies, D., Millan, F., Retterer, K., Dyack, S., MacKay, S., Morales-Gonzalez, S., Giunta, M., Munro, B., Hudson, G., Scavina, M., Baker, L., Massini, T. C., Lek, M., Hu, Y., Ezzo, D., AlKuraya, F. S., Kang, P. B., Griffin, H., Foley, A. R., Schuelke, M., Horvath, R., and Bonnemann, C. G. (2018) Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy. *Am J Hum Genet* **102**, 858-873

11. Gillespie, A., Gabunilas, J., Jen, J. C., and Chanfreau, G. F. (2017) Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in *S. cerevisiae*. *RNA* **23**, 466-472
12. Fasken, M. B., Losh, J. S., Leung, S. W., Brutus, S., Avin, B., Vaught, J. C., Potter-Birriel, J., Craig, T., Conn, G. L., Mills-Lujan, K., Corbett, A. H., and van Hoof, A. (2017) Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast. *Genetics* **205**, 221-237
13. Morton, D. J., Kuiper, E. G., Jones, S. K., Leung, S. W., Corbett, A. H., and Fasken, M. B. (2018) The RNA exosome and RNA exosome-linked disease. *RNA* **24**, 127-142
14. Morton, D. J., Jalloh, B., Kim, L., Kremsky, I., Le, T., Nguyen, K. B., Rounds, J. C., Sterrett, M. C., Brown, B., Leung, S. W., Fasken, M. B., Moberg, K. H., and Corbett, A. H. (*In press*, 2020) A *Drosophila* Model of Pontocerebellar Hypoplasia Reveals a Critical Role for the RNA Exosome in Neurons. *PLOS Genetics*
15. Yang, X., Bayat, V., DiDonato, N., Zhao, Y., Zarnegar, B., Siprashvili, Z., Lopez-Pajares, V., Sun, T., Tao, S., Li, C., Rump, A., Khavari, P., and Lu, B. (2019) Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis. *Hum Mol Genet*
16. Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schoneborn, S., Menezes, M. P., Hong, J. E., Leong, D. W., Senderek, J., Salman, M. S., Chitayat, D., Seeman, P., von Moers, A., Graul-Neumann, L., Kornberg, A. J., Castro-Gago, M., Sobrido, M. J., Sanefuji, M., Shieh, P. B., Salamon, N., Kim, R. C., Vinters, H. V., Chen, Z., Zerres, K., Ryan, M. M., Nelson, S. F., and Jen, J. C. (2012) Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat Genet* **44**, 704-708
17. Mitchell, P., Petfalski, E., and Tollervey, D. (1996) The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. *Genes & Development* **10**, 502-513
18. Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J* **18**, 5399-5410
19. Liu, Q., Greimann, J. C., and Lima, C. D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**, 1223-1237
20. Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G., and Conti, E. (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nature structural & molecular biology* **12**, 575-581
21. Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell* **139**, 547-559
22. Januszyk, K., and Lima, C. D. (2010) Structural components and architectures of RNA exosomes. *Advances in experimental medicine and biology* **702**, 9-28
23. Lorentzen, E., Basquin, J., and Conti, E. (2008) Structural organization of the RNA-degrading exosome. *Current opinion in structural biology* **18**, 709-713
24. Delan-Forino, C., Schneider, C., and Tollervey, D. (2017) Transcriptome-wide analysis of alternative routes for RNA substrates into the exosome complex. *Plos Genetics* **13**, 26
25. Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012) Transcriptome-wide analysis of exosome targets. *Mol Cell* **48**, 422-433

26. Wu, G., Schmid, M., Rib, L., Polak, P., Meola, N., Sandelin, A., and Jensen, T. H. (2020) A Two-Layered Targeting Mechanism Underlies Nuclear RNA Sorting by the Human Exosome. *Cell Rep* **30**, 2387-2401.e2385
27. Schmid, M., and Jensen, T. H. (2019) The Nuclear RNA Exosome and Its Cofactors. *Advances in experimental medicine and biology* **1203**, 113-132
28. LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**, 713-724
29. Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M. S., Mapendano, C. K., Schierup, M. H., and Jensen, T. H. (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851-1854
30. Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., Régnauld, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., and Jacquier, A. (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725-737
31. Klauer, A. A., and van Hoof, A. (2012) Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. *Wiley Interdiscip Rev RNA* **3**, 649-660
32. Kögel, A., Keidel, A., Bonneau, F., Schäfer, I. B., and Conti, E. (2022) The human SKI complex regulates channeling of ribosome-bound RNA to the exosome via an intrinsic gatekeeping mechanism. *Mol Cell* **82**, 756-769.e758
33. Zinder, J. C., and Lima, C. D. (2017) Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Gene & Development* **2**, 88-100
34. Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Raijmakers, R., Pruijn, G. J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107**, 451-464
35. Schilders, G., Raijmakers, R., Raats, J. M., and Pruijn, G. J. (2005) MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. *Nucleic acids research* **33**, 6795-6804
36. Puno, M. R., and Lima, C. D. (2018) Structural basis for MTR4-ZCCHC8 interactions that stimulate the MTR4 helicase in the nuclear exosome-targeting complex. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5506-E5515
37. Kowalinski, E., Kogel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016) Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. *Mol. Cell* **63**, 125-134
38. Lubas, M., Christensen, M. S., Kristiansen, M. S., Domanski, M., Falkenby, L. G., Lykke-Andersen, S., Andersen, J. S., Dziembowski, A., and Jensen, T. H. (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Mol Cell* **43**, 624-637
39. Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J. S., Sandelin, A., and Jensen, T. H. (2016) Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. *Mol Cell* **64**, 520-533
40. Dziembowski, A., Lorentzen, E., Conti, E., and Seraphin, B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nature structural & molecular biology* **14**, 15-22

41. Somashekar, P. H., Kaur, P., Stephen, J., Guleria, V. S., Kadavigere, R., Girisha, K. M., Bielas, S., Upadhyai, P., and Shukla, A. (2021) Bi-allelic missense variant, p. Ser35Leu in EXOSC1 is associated with pontocerebellar hypoplasia. *Clinical Genetics* **99**, 594-600
42. Di Donato, N., Neuhann, T., Kahlert, A. K., Klink, B., Hackmann, K., Neuhann, I., Novotna, B., Schallner, J., Krause, C., Glass, I. A., Parnell, S. E., Benet-Pages, A., Nissen, A. M., Berger, W., Altmuller, J., Thiele, H., Weber, B. H., Schrock, E., Dobyns, W. B., Bier, A., and Rump, A. (2016) Mutations in EXOSC2 are associated with a novel syndrome characterised by retinitis pigmentosa, progressive hearing loss, premature ageing, short stature, mild intellectual disability and distinctive gestalt. *J Med Genet* **53**, 419-425
43. Slavotinek, A., Misceo, D., Htun, S., Mathisen, L., Frengen, E., Foreman, M., Hurtig, J. E., Enyenihi, L., Sterrett, M. C., and Leung, S. W. (2020) Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness. *Human Molecular Genetics* **29**, 2218-2239
44. Fasken, M. B., Morton, D. J., Kuiper, E. G., Jones, S. K., Leung, S. W., and Corbett, A. H. (2020) The RNA Exosome and Human Disease. *Methods in molecular biology (Clifton, N.J.)* **2062**, 3-33
45. Rudnik-Schoneborn, S., Senderek, J., Jen, J. C., Houge, G., Seeman, P., Puchmajerova, A., Graul-Neumann, L., Seidel, U., Korinthenberg, R., Kirschner, J., Seeger, J., Ryan, M. M., Muntoni, F., Steinlin, M., Sztriha, L., Colomer, J., Hubner, C., Brockmann, K., Van Maldergem, L., Schiff, M., Holzinger, A., Barth, P., Reardon, W., Yourshaw, M., Nelson, S. F., Eggermann, T., and Zerres, K. (2013) Pontocerebellar hypoplasia type 1: clinical spectrum and relevance of EXOSC3 mutations. *Neurology* **80**, 438-446
46. Bizzari, S., Hamzeh, A. R., Mohamed, M., Al-Ali, M. T., and Bastaki, F. (2019) Expanded PCH1D phenotype linked to EXOSC9 mutation. *European journal of medical genetics*, 103622
47. Slavotinek, A., Misceo, D., Htun, S., Mathisen, L., Frengen, E., Foreman, M., Hurtig, J. E., Enyenihi, L., Sterrett, M. C., Leung, S. W., Schneidman-Duhovny, D., Estrada-Veras, J., Duncan, J. L., Xia, V., Beleford, D., Si, Y., Douglas, G., Treidene, H. E., van Hoof, A., Fasken, M. B., and Corbett, A. H. (In press, 2020) Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness. *Human Molecular Genetics*
48. Yang, X., Bayat, V., DiDonato, N., Zhao, Y., Zarnegar, B., Siprashvili, Z., Lopez-Pajares, V., Sun, T., Tao, S., Li, C., Rump, A., Khavari, P., and Lu, B. (2019) Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis. *Human molecular genetics* **29**, 541-553
49. Kiss, D. L., and Andrusis, E. D. (2011) The exozyme model: a continuum of functionally distinct complexes. *Rna* **17**, 1-13
50. Lim, S. J., Boyle, P. J., Chinen, M., Dale, R. K., and Lei, E. P. (2013) Genome-wide localization of exosome components to active promoters and chromatin insulators in *Drosophila*. *Nucleic acids research* **41**, 2963-2980
51. Bizzari, S., Hamzeh, A. R., Mohamed, M., Al-Ali, M. T., and Bastaki, F. (2020) Expanded PCH1D phenotype linked to EXOSC9 mutation. *European journal of medical genetics* **63**, 103622

52. Akalal, D. B., Wilson, C. F., Zong, L., Tanaka, N. K., Ito, K., and Davis, R. L. (2006) Roles for *Drosophila* mushroom body neurons in olfactory learning and memory. *Learning & memory (Cold Spring Harbor, N.Y.)* **13**, 659-668
53. Goldowitz, D., and Hamre, K. (1998) The cells and molecules that make a cerebellum. *Trends Neurosci* **21**, 375-382
54. Bellen, H. J., Wangler, M. F., and Yamamoto, S. (2019) The fruit fly at the interface of diagnosis and pathogenic mechanisms of rare and common human diseases. *Hum Mol Genet* **28**, R207-R214
55. Bates, A. S., Janssens, J., Jefferis, G. S., and Aerts, S. (2019) Neuronal cell types in the fly: single-cell anatomy meets single-cell genomics. *Current opinion in neurobiology* **56**, 125-134
56. Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A. N., and Basu, U. (2014) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* **514**, 389-393
57. Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J., and Slonimski, P. P. (1989) Birth of the D-E-A-D box. *Nature* **337**, 121-122
58. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015) GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Human mutation* **36**, 928-930
59. Januszyk, K., and Lima, C. D. (2014) The eukaryotic RNA exosome. *Current opinion in structural biology* **24**, 132-140
60. Kadowaki, T., Schneider, R., Hitomi, M., and Tartakoff, A. M. (1995) Mutations in nucleolar proteins lead to nucleolar accumulation of polyA⁺ RNA in *Saccharomyces cerevisiae*. *Mol Biol Cell* **6**, 1103-1110
61. Hou, D., Ruiz, M., and Andrulis, E. D. (2012) The ribonuclease Dis3 is an essential regulator of the developmental transcriptome. *BMC Genomics* **13**, 359
62. Morton, D. J., Jalloh, B., Kim, L., Kremisky, I., Nair, R. J., Nguyen, K. B., Rounds, J. C., Sterrett, M. C., Brown, B., and Le, T. (2020) A *Drosophila* model of Pontocerebellar Hypoplasia reveals a critical role for the RNA exosome in neurons. *PLoS Genetics* **16**, e1008901
63. de Amorim, J., Slavotinek, A., Fasken, M. B., Corbett, A. H., and Morton, D. J. (2020) Modeling Pathogenic Variants in the RNA Exosome. *RNA Dis* **7**
64. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyrto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Pontén, F. (2015) Tissue-based map of the human proteome. *Science* **347**, 1260419
65. Slavotinek, A., Misceo, D., Htun, S., Mathisen, L., Frengen, E., Foreman, M., Hurtig, J. E., Enyenihi, L., Sterrett, M. C., Leung, S. W., Schneidman-Duhovny, D., Estrada-Veras, J., Duncan, J. L., Haaxma, C. A., Kamsteeg, E. J., Xia, V., Belefors, D., Si, Y., Douglas, G., Treidene, H. E., van Hoof, A., Fasken, M. B., and Corbett, A. H. (2020) Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness. *Hum Mol Genet* **29**, 2218-2239

66. de Amorim, J. L., Leung, S. W., Haji-Seyed-Javadi, R., Hou, Y., Yu, D. S., Ghalei, H., Khoshnevis, S., Yao, B., and Corbett, A. H. (2023) The RNA helicase DDX1 associates with the nuclear RNA exosome and modulates R-loops. *bioRxiv*, 2023.2004. 2017.537228
67. Houseley, J., LaCava, J., and Tollervey, D. (2006) RNA-quality control by the exosome. *Nature reviews. Molecular cell biology* **7**, 529-539
68. Klauer, A. A., and van Hoof, A. (2012) Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. *Wiley interdisciplinary reviews. RNA* **3**, 649-660
69. Rigby, R. E., and Rehwinkel, J. (2015) RNA degradation in antiviral immunity and autoimmunity. *Trends Immunol* **36**, 179-188
70. Molleston, J. M., Sabin, L. R., Moy, R. H., Menghani, S. V., Rausch, K., Gordesky-Gold, B., Hopkins, K. C., Zhou, R., Jensen, T. H., Wilusz, J. E., and Cherry, S. (2016) A conserved virus-induced cytoplasmic TRAMP-like complex recruits the exosome to target viral RNA for degradation. *Genes & Development* **30**, 1658-1670
71. Schneider, C., and Tollervey, D. (2013) Threading the barrel of the RNA exosome. *Trends in biochemical sciences* **38**, 485-493
72. Wasmuth, E. V., and Lima, C. D. (2012) Structure and Activities of the Eukaryotic RNA Exosome. *The Enzymes* **31**, 53-75
73. Wasmuth, E. V., Januszyk, K., and Lima, C. D. (2014) Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature* **511**, 435-439
74. Wasmuth, E. V., Zinder, J. C., Zattas, D., Das, M., and Lima, C. D. (2017) Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase. *eLife* **6**
75. Damseh, N. S., Obeidat, A. N., Ahammed, K. S., Al-Ashhab, M., Awad, M. A., and van Hoof, A. (2023) Pontocerebellar hypoplasia associated with p. Arg183Trp homozygous variant in EXOSC1 gene: A case report. *American Journal of Medical Genetics Part A*
76. Spyridakis, A. C., Cao, Y., and Litra, F. (2022) A Rare Case of Pontocerebellar Hypoplasia Type 1B With Literature Review. *Cureus* **14**, e27098
77. Ohguchi, Y., and Ohguchi, H. (2023) DIS3: The Enigmatic Gene in Multiple Myeloma. *Int J Mol Sci* **24**
78. Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**, 227-232
79. Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schoneborn, S., Menezes, M. P., Hong, J. E., Leong, D. W., Senderek, J., Salman, M. S., Chitayat, D., Seeman, P., von Moers, A., Graul-Neumann, L., Kornberg, A. J., Castro-Gago, M., Sobrido, M. J., Sanefuji, M., Shieh, P. B., Salamon, N., Kim, R. C., Vinters, H. V., Chen, Z., Zerres, K., Ryan, M. M., Nelson, S. F., and Jen, J. C. (2012) Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat Genet* **44**, 704-708
80. Tomecki, R., Kristiansen, M. S., Lykke-Andersen, S., Chlebowski, A., Larsen, K. M., Szczesny, R. J., Drazkowska, K., Pastula, A., Andersen, J. S., Stepień, P. P., Dziembowski, A., and Jensen, T. H. (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *Embo j* **29**, 2342-2357
81. Makino, D. L., Schuch, B., Stegmann, E., Baumgartner, M., Basquin, C., and Conti, E. (2015) RNA degradation paths in a 12-subunit nuclear exosome complex. *Nature* **524**, 54-58

82. Schaeffer, D., Tsanova, B., Barbas, A., Reis, F. P., Dastidar, E. G., Sanchez-Rotunno, M., Arraiano, C. M., and van Hoof, A. (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature structural & molecular biology* **16**, 56-62
83. Schuller, J. M., Falk, S., Fromm, L., Hurt, E., and Conti, E. (2018) Structure of the nuclear exosome captured on a maturing preribosome. *Science* **360**, 219-222
84. Fan, J., Kuai, B., Wu, G. F., Wu, X. D., Chi, B. K., Wang, L. T., Wang, K., Shi, Z. B., Zhang, H., Chen, S., He, Z. S., Wang, S. Y., Zhou, Z. C., Li, G. H., and Cheng, H. (2017) Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *Embo J.* **36**, 2870-2886
85. Lejeune, F., Li, X., and Maquat, L. E. (2003) Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell* **12**, 675-687
86. van Hoof, A., Staples, R. R., Baker, R. E., and Parker, R. (2000) Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol Cell Biol* **20**, 8230-8243
87. Marin-Vicente, C., Domingo-Prim, J., Eberle, A. B., and Visa, N. (2015) RRP6/EXOSC10 is required for the repair of DNA double-strand breaks by homologous recombination. *Journal of cell science* **128**, 1097-1107
88. Richard, P., and Manley, J. L. (2017) R loops and links to human disease. *Journal of molecular biology* **429**, 3168-3180
89. Richard, P., Feng, S., and Manley, J. L. (2013) A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes & Development* **27**, 2227-2232
90. Namavar, Y., Barth, P. G., Poll-The, B. T., and Baas, F. (2011) Classification, diagnosis and potential mechanisms in pontocerebellar hypoplasia. *Orphanet journal of rare diseases* **6**, 50
91. Ivanov, I., Atkinson, D., Litvinenko, I., Angelova, L., Andonova, S., Mumdjiev, H., Pacheva, I., Panova, M., Yordanova, R., Belovejdov, V., Petrova, A., Bosheva, M., Shmilev, T., Savov, A., and Jordanova, A. (2018) Pontocerebellar hypoplasia type 1 for the neuropediatrician: Genotype-phenotype correlations and diagnostic guidelines based on new cases and overview of the literature. *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society* **22**, 674-681
92. Radvanska, E., Pos, Z., Zatkova, A., Hyblova, M., Bauer, F., Szemes, T., Kadasi, L., and Radvanszky, J. (2022) Molecularly confirmed pontocerebellar hypoplasia in a large family from Slovakia with four severely affected children. *Bratisl Lek Listy* **123**, 568-572
93. Halbach, F., Rode, M., and Conti, E. (2012) The crystal structure of *S. cerevisiae* Ski2, a DExH helicase associated with the cytoplasmic functions of the exosome. *Rna* **18**, 124-134
94. Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019) A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. *Embo J.* **38**
95. Makino, D. L., Baumgärtner, M., and Conti, E. (2013) Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* **495**, 70-75

96. Schmid, S. R., and Linder, P. (1992) D-E-A-D protein family of putative RNA helicases. *Mol Microbiol* **6**, 283-291
97. Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J., and Slonimski, P. P. (1989) Birth of the DEAD box. *Nature* **337**, 121-122
98. Godbout, R., Hale, M., and Bisgrove, D. (1994) A human DEAD box protein with partial homology to heterogeneous nuclear ribonucleoprotein U. *Gene* **138**, 243-245
99. Suzuki, T., Katada, E., Mizuoka, Y., Takagi, S., Kazuki, Y., Oshimura, M., Shindo, M., and Hara, T. (2021) A novel all-in-one conditional knockout system uncovered an essential role of DDX1 in ribosomal RNA processing. *Nucleic acids research* **49**, e40
100. Ribeiro de Almeida, C., Dhir, S., Dhir, A., Moghaddam, A. E., Sattentau, Q., Meinhart, A., and Proudfoot, N. J. (2018) RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol Cell* **70**, 650-662.e658
101. Li, L., Germain, D. R., Poon, H. Y., Hildebrandt, M. R., Monckton, E. A., McDonald, D., Hendzel, M. J., and Godbout, R. (2016) DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol Cell Biol* **36**, 2794-2810
102. Li, L., Monckton, E. A., and Godbout, R. (2008) A role for DEAD box 1 at DNA double-strand breaks. *Mol Cell Biol* **28**, 6413-6425
103. Passinen, S., Valkila, J., Manninen, T., Syvala, H., and Ylikomi, T. (2001) The C-terminal half of Hsp90 is responsible for its cytoplasmic localization. *European Journal of Biochemistry* **268**, 5337-5342
104. Ye, K. Q. (2005) Nucleophosmin/B23, a multifunctional protein that can regulate apoptosis. *Cancer Biology & Therapy* **4**, 918-923
105. Perez-Gonzalez, A., Pazo, A., Navajas, R., Ciordia, S., Rodriguez-Frandsen, A., and Nieto, A. (2014) hCLE/C14orf166 associates with DDX1-HSPC117-FAM98B in a novel transcription-dependent shuttling RNA-transporting complex. *PloS one* **9**, e90957
106. Godbout, R., Packer, M., and Bie, W. (1998) Overexpression of a DEAD box protein (DDX1) in neuroblastoma and retinoblastoma cell lines. *Journal of Biological Chemistry* **273**, 21161-21168
107. Chen, H. C., Lin, W. C., Tsay, Y. G., Lee, S. C., and Chang, C. J. (2002) An RNA helicase, DDX1, interacting with poly(A) RNA and heterogeneous nuclear ribonucleoprotein K. *J Biol Chem* **277**, 40403-40409
108. Ryan, A. J., Squires, S., Strutt, H. L., and Johnson, R. T. (1991) Camptothecin cytotoxicity in mammalian cells is associated with the induction of persistent double strand breaks in replicating DNA. *Nucleic acids research* **19**, 3295-3300
109. Kuo, L. J., and Yang, L.-X. (2008) γ -H2AX-a novel biomarker for DNA double-strand breaks. *In vivo* **22**, 305-309
110. Pirouz, M., Munafò, M., Ebrahimi, A. G., Choe, J., and Gregory, R. I. (2019) Exonuclease requirements for mammalian ribosomal RNA biogenesis and surveillance. *Nature structural & molecular biology* **26**, 490-500
111. Cargill, M., Venkataraman, R., and Lee, S. (2021) DEAD-Box RNA Helicases and Genome Stability. *Genes (Basel)* **12**
112. Suzuki, T., Katada, E., Mizuoka, Y., Takagi, S., Kazuki, Y., Oshimura, M., Shindo, M., and Hara, T. (2021) A novel all-in-one conditional knockout system uncovered an

- essential role of DDX1 in ribosomal RNA processing. *Nucleic Acids Research* **49**, e40-e40
113. Henras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A., and Gleizes, P. E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley interdisciplinary reviews. RNA* **6**, 225-242
 114. Burman, L. G., and Mauro, V. P. (2012) Analysis of rRNA processing and translation in mammalian cells using a synthetic 18S rRNA expression system. *Nucleic acids research* **40**, 8085-8098
 115. Leung, E., and Brown, J. D. (2010) Biogenesis of the signal recognition particle. *Biochemical Society Transactions* **38**, 1093-1098
 116. Wang, I. X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M., and Cheung, V. G. (2018) Human proteins that interact with RNA/DNA hybrids. *Genome research* **28**, 1405-1414
 117. Laffleur, B., Lim, J., Zhang, W., Chen, Y., Pefanis, E., Bizarro, J., Batista, C. R., Wu, L., Economides, A. N., and Wang, J. (2021) Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nature genetics* **53**, 230-242
 118. Zhang, M., Chen, D., Xia, J., Han, W., Cui, X., Neuenkirchen, N., Hermes, G., Sestan, N., and Lin, H. (2017) Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes & Development* **31**, 1354-1369
 119. Telekawa, C., Boisvert, F. M., and Bachand, F. (2018) Proteomic profiling and functional characterization of post-translational modifications of the fission yeast RNA exosome. *Nucleic acids research* **46**, 11169-11183
 120. Crossley, M. P., Bocek, M., and Cimprich, K. A. (2019) R-loops as cellular regulators and genomic threats. *Mol. Cell* **73**, 398-411
 121. Nair, L., Chung, H., and Basu, U. (2020) Regulation of long non-coding RNAs and genome dynamics by the RNA surveillance machinery. *Nature reviews Molecular cell biology* **21**, 123-136
 122. Nojima, T., and Proudfoot, N. J. (2022) Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. *Nature Reviews Molecular Cell Biology* **23**, 389-406
 123. Tatomer, D. C., Elrod, N. D., Liang, D., Xiao, M.-S., Jiang, J. Z., Jonathan, M., Huang, K.-L., Wagner, E. J., Cherry, S., and Wilusz, J. E. (2019) The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes & Development* **33**, 1525-1538
 124. Zhang, F., Ma, T., and Yu, X. (2013) A core hSSB1-INTS complex participates in the DNA damage response. *Journal of cell science* **126**, 4850-4855
 125. Wagnon, J. L., Briese, M., Sun, W., Mahaffey, C. L., Curk, T., Rot, G., Ule, J., and Frankel, W. N. (2012) CELF4 regulates translation and local abundance of a vast set of mRNAs, including genes associated with regulation of synaptic function. *PLoS genetics* **8**, e1003067
 126. Halgren, C., Bache, I., Bak, M., Myatt, M. W., Anderson, C. M., Brøndum-Nielsen, K., and Tommerup, N. (2012) Haploinsufficiency of CELF4 at 18q12. 2 is associated with developmental and behavioral disorders, seizures, eye manifestations, and obesity. *European Journal of Human Genetics* **20**, 1315-1319
 127. Yang, Y., Mahaffey, C. L., Bérubé, N., Maddatu, T. P., Cox, G. A., and Frankel, W. N. (2007) Complex seizure disorder caused by Brunol4 deficiency in mice. *PLoS genetics* **3**, e124

128. Neasta, J., Uttenweiler-Joseph, S., Chaoui, K., Monsarrat, B., Meunier, J. C., and Moulédous, L. (2006) Effect of long-term exposure of SH-SY5Y cells to morphine: a whole cell proteomic analysis. *Proteome Sci* **4**, 23
129. Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A. N., and Basu, U. (2014) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* **514**, 389-393
130. Zhao, Z., Chen, X., Dowbaj, A. M., Sljukic, A., Bratlie, K., Lin, L., Fong, E. L. S., Balachander, G. M., Chen, Z., Soragni, A., Huch, M., Zeng, Y. A., Wang, Q., and Yu, H. (2022) Organoids. *Nature Reviews Methods Primers* **2**, 94
131. Dutschmann, M., and Dick, T. E. (2012) Pontine mechanisms of respiratory control. *Compr Physiol* **2**, 2443-2469
132. Colombari, E., Sato, M. A., Cravo, S. L., Bergamaschi, C. T., Campos Jr, R. R., and Lopes, O. U. (2001) Role of the medulla oblongata in hypertension. *Hypertension* **38**, 549-554
133. Strick, P. L., Dum, R. P., and Fiez, J. A. (2009) Cerebellum and nonmotor function. *Annual review of neuroscience* **32**, 413-434
134. Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature reviews genetics* **13**, 227-232
135. Sterrett, M. C., Farchi, D., Strassler, S. E., Boise, L. H., Fasken, M. B., and Corbett, A. H. (2023) In vivo characterization of the critical interaction between the RNA exosome and the essential RNA helicase Mtr4 in *Saccharomyces cerevisiae*. *G3 Genes|Genomes|Genetics*
136. Lingaraju, M., Johnsen, D., Schlundt, A., Langer, L. M., Basquin, J., Sattler, M., Heick Jensen, T., Falk, S., and Conti, E. (2019) The MTR4 helicase recruits nuclear adaptors of the human RNA exosome using distinct arch-interacting motifs. *Nature communications* **10**, 3393
137. Weick, E. M., Puno, M. R., Januszyk, K., Zinder, J. C., DiMattia, M. A., and Lima, C. D. (2018) Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex. *Cell* **173**, 1663-1677
138. Godbout, R., Li, L., Liu, R.-Z., and Roy, K. (2007) Role of DEAD box 1 in retinoblastoma and neuroblastoma.