

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:

_____ Date

**Characterizing Human-Disease Mutations in the Essential
RNA Exosome with the Budding Yeast Model System**

By

Maria C. Sterrett

B.A. Biology, Carleton College, 2014

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

Anita H. Corbett
Advisor

Victor Faundez
Committee Member

Yue Feng
Committee Member

Ken Moberg
Committee Member

Bing Yao
Committee Member

Accepted:

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

Date

Characterizing Human-Disease Mutations in the Essential RNA Exosome with the Budding Yeast Model System

By

Maria C. Sterrett

B.A. Biology, Carleton College, 2014

Advisor: Anita 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 Biological and Biomedical Sciences
2023

Abstract

Characterizing Human-Disease Mutations in the Essential RNA Exosome with the Budding Yeast Model System

By Maria C. Sterrett

Differential genetic expression allows for the cellular diversity and specialized tissues composing our bodies. Key to this specialization are post-transcriptional events that regulate RNA. The RNA exosome is critical for many post-transcriptional events, required for 3' to 5' processing and degradation of a vast amount of RNA. This molecular machine is highly conserved, both in function and in structure, consisting of a 3-subunit cap, a 6-subunit core ring and a catalytic 3'-5' exo/endoribonuclease. Recently, missense mutations have been identified in the structural cap and core subunit genes that cause distinct disease pathologies, including neurological and developmental disorders. The identification of these distinct diseases raises the question of how single amino acid substitutions in conserved, structural subunits of this complex affect RNA exosome function and consequently underlie diverse pathologies.

To explore the functional consequences of disease-linked amino acid substitutions in RNA exosome structural subunits, we generated *Saccharomyces cerevisiae* modeling disease mutations in corresponding budding yeast genes. In **Chapter II**, I present a systematic characterization of a single *S. cerevisiae* RNA exosome mutant model, *rrp4-G226D*. The *rrp4-G226D* cells model a mutation found in the cap subunit gene *EXOSC2* that is linked to a novel syndrome. I uncover that the modeled pathogenic amino acid substitution impacts the function of the RNA exosome through disruptions of a key interaction with an RNA helicase known as Mtr4. In **Chapter III**, I take similar methods as in Chapter II and characterize a the *rrp4-M68T* mutant model. The *rrp4-M68T* cells model a mutation in *EXOSC2* that is linked to multiple-myeloma. I find that *rrp4-M68T* also destabilizes interactions between the RNA exosome and Mtr4, however *rrp4-M68T* cells have functional consequences that differ from those of the *rrp4-G226D* model. In **Chapter IV**, I present a comparative RNA-Seq analysis of several *S. cerevisiae* models of disease-linked mutations in RNA exosome structural genes. I identify broad transcriptomic changes and uncover a shared link between RNA exosome disease mutations and translational defects. Overall, these studies provide evidence that disease-linked amino acid substitutions in structural subunits of the RNA exosome impair the function of the essential complex through differential *in vivo* consequences.

Characterizing Human-Disease Mutations in the Essential RNA Exosome with the Budding Yeast Model System

By

Maria C. Sterrett

B.A. Biology, Carleton College, 2014

Advisor: Anita 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 Biological and Biomedical Sciences
2023

Acknowledgements

“Love is a better teacher than a sense of duty” -Albert Einstein

In thinking about these acknowledgements, I am drawn to this quote by one of the greatest minds of modern time, as I interpret it to have two meanings:

- 1) That *passion* will get you further in achieving your goals.
- 2) That *passionate mentors* will impact you the most.

I would like to express my deepest gratitude to all the passionate mentors who taught me how to pursue academic research and education with creativity, curiosity, and, most of all, “love”.

First and foremost, that includes my PI, Dr. Anita Corbett. Dr. Corbett showed me how passionate mentorship and academic research can go hand-in-hand and create a thriving learning environment that grows student curiosity while advancing scientific discovery. I thank her for the continual support and encouragement as I pursued teaching and mentorship opportunities in addition to my doctoral studies. Dr. Corbett’s guidance over my training and her influence has been and will be a tangible force that shapes me as a scientist and educator. I am forever grateful for her mentorship and friendship.

Second, I would like to thank the amazing, talented, and passionate people in the Corbett lab, past and present—Dr. Laramie Lemon, Dr. Derrick Morton, Julia de Amorim, Carly Lancaster, Jordan Goldy, Dr. Celina Jones, Dr. Milo Fasken, Dr. Sara Leung—who I count as dear and treasured friends. Dr. Milo Fasken and Dr. Sara Leung truly served as second advisors to me as I conducted my doctoral research. I thank them both for helping me through experimental successes and setbacks, always being an open ear for brainstorming and analysis of new results and failures. I very much appreciate learning from their expertise and witnessing their passion in the pursuit of scientific discovery enhance not only my research but that of the entire lab. Much of the research presented here is built upon foundational work done by Dr. Fasken. I am particularly thankful to him for the continual support and guidance as I pursued my thesis. I am also thankful for my “lab twin” Julia de Amorim. Julia and I joined Dr. Corbett’s lab group at the same time and have been at each other’s side ever since. I am extremely fortunate to have had a support system like Julia and I thank her for being a constant throughout my time as a Corbett lab member.

Third, I’d like to thank my Emory scientific community. This includes my committee members for their feedback, encouragement, and guidance throughout my graduate experience. Thank you to Drs. Ken Moberg, Yue Feng, Bing Yao and Victor Faundez. They each brought their own curiosity and creativity to my research project, further infusing my Emory experience with passion and excitement. I also want to thank my entire BCDB community. I am very humbled and grateful to have found a community that embraced me and nurtured me as I pursued my own passions in research and education. Thank you to the BCDB faculty for providing space for students to grow both as scientists and impassioned people. And thank you to my BCDB peers who have inspired me to continue in academia with joy, enthusiasm, and compassion. The friendships I have gained are enduring and will bolster me throughout my lifetime.

Fourth, I’d like to thank the numerous student mentees I had the privilege of working with at Emory. Jae Bucknor, Lakshya Seth, Shelly Seth, Samika Joshi, Liz Enyenihi, Daniela Farchi, Robin Mo, Will Ball, Jen Dean, Foje-Geh Tendoh—thank you for trusting me and for bringing your excitement, your enthusiasm, your passions, and your identities into the lab. The experiences we shared not only enhanced my doctoral research but also helped me grow as a mentor and teacher.

Lastly, I’d like to thank the people who truly exemplify “love” in my life. Thank you, Mom and Dad, for always supporting your independent little girl in pursuing her curiosity. Thank you Christian for believing in me as my best friend and partner. And thank you Janet, Craig, Ellie, Ryan, Caroline, and Drew for making Atlanta my home. Without you, I would not have achieved all that I have, nor would I be all that I am.

Table of Contents

Chapter I: Introduction.....	1
1.1 Post-Transcriptional Regulation of Gene Expression	2
1.2 RNA Turnover and Decay	5
1.3 The RNA Exosome: An Essential Molecular Machine	7
1.4 Function of the Eukaryotic RNA Exosome	9
1.5 Interacting Cofactors Expand the Function of the RNA Exosome	10
1.6 RNA Exosome and Human Disease: RNA Exosomopathies	11
1.7 RNA Exosomopathy-linked Amino Acid Substitutions and Differential Effects on RNA exosome Function <i>in vivo</i>	13
1.8 Doctoral Research Aims	15
1.9 Chapter I Figures.....	17
Figure 1. Schematic representation of mRNA post-transcriptional regulation and gene expression..	17
Figure 2. Conservation of the core RNA exosome structure in eukarya and archaea.....	20
Figure 3. Overview of pathogenic RNA exosomopathy amino acid substitutions in the human RNA exosome.	23
Figure 4. Proposed model of how RNA exosomopathy amino acid substitutions may result in distinct disease pathologies.....	25
1.10 References.....	27
Chapter II: A Budding Yeast Model for Human Disease Mutations in the EXOSC2 Cap Subunit of the RNA Exosome Complex	35
2.1 Abstract.....	36
2.2 Introduction.....	37
2.3 Materials and Methods.....	41
2.4 Results.....	51
2.5 Discussion.....	61
2.6 Chapter II Acknowledgements.....	67
2.7 Chapter II Figures	68
Figure 1. Overview of pathogenic amino acid substitutions in the human cap subunit EXOSC2 of the RNA exosome.....	68
Figure 2. Modeling pathogenic amino acid substitutions in Human EXOSC2 and <i>S. cerevisiae</i> Rrp4.	70
Figure 3. <i>S. cerevisiae</i> Rrp4 variants that model EXOSC2 variants identified in patients show impaired function.	72

Figure 4. The Rrp4 G58V and Rrp4 G226D variants do not associate with the RNA exosome complex in the presence of wild-type Rrp4, but the Rrp4 G226D variant can associate with the RNA exosome complex.....	74
Figure 5. The <i>rrp4-G226D</i> variant cells show elevated levels of some but not all RNA exosome target transcripts.....	76
Figure 6. RNA-Seq analysis of <i>rrp4-G226D</i> cells reveal distinct transcriptomic changes compared to <i>RRP4</i> cells.....	79
Figure 7. Validation of the differentially expressed transcripts identified in the RNA-Seq confirms that the levels of key mRNAs and CUTs are significantly altered in <i>rrp4-G226D</i> cells and reveals that some of these transcripts are not changed in <i>rrp40-W195R</i> cells.....	81
Figure 8. The <i>rrp4-G226D</i> mutant exhibits distinct negative genetic interactions with RNA exosome cofactor mutants that are not shared by the <i>rrp40-W195R</i> mutant.....	82
Figure 9. The <i>rrp4-G226D</i> mutant shows genetic interaction with an <i>mtr4</i> mutant that is impaired for interaction with Rrp6/Rrp47 and Rrp4 G226D impairs interaction with Mtr4.....	84
2.8 Chapter II Supplementary Materials.....	86
Supplementary Figure S1. Protein sequence alignment of human EXOSC2 and EXOSC3.....	86
Supplementary Figure S2. Modeling of the human EXOSC2-EXOSC4 and yeast Rrp4-Rrp41 interface show structural conservation.....	88
Supplementary Figure S3. Increased input signal levels for rRNA northern blot in Figure 5A emphasize previously observed accumulation of 5.8S precursors in <i>rrp40-W195R</i> cells.....	89
Supplementary Figure S4. Volcano plot of autophagy transcripts differentially expressed in the <i>rrp4-G226D</i> RNA-Seq.....	90
Table S1. <i>S. cerevisiae</i> Strains and Plasmids used in this study.....	91
Table S2. DNA Oligonucleotides employed for RT-qPCR.....	92
Table S3. Summary of in silico predictions for pathogenic amino acid substitutions in EXOSC2 and Rrp4.....	93
2.9 References.....	94
Chapter III: <i>In vivo</i> Characterization of the Critical Interaction between the RNA Exosome and the Essential RNA Helicase Mtr4 in <i>Saccharomyces cerevisiae</i>.....	100
3.1 Abstract.....	101
3.2 Introduction.....	102
3.3 Materials and Methods.....	107
3.4 Results.....	117
3.5 Discussion.....	130
3.6 Chapter III Acknowledgements.....	136
3.7 Chapter III Figures.....	136
Figure 1. Overview of multiple myeloma linked amino acid substitutions in the human cap subunit EXOSC2 of the RNA exosome.....	137

Figure 2. Modeling the multiple myeloma EXOSC2 M40T amino acid substitution in the human EXOSC2 cap subunit and the <i>S. cerevisiae</i> ortholog Rrp4.....	139
Figure 3. <i>S. cerevisiae rrp4-M68T</i> mutant cells that model the EXOSC2 M40T variant identified in multiple myeloma patients show impaired function on drugs that impact RNA processing.....	142
Figure 4. The <i>rrp4-M68T</i> mutant cells show elevated levels of specific RNA exosome target transcripts that depend on the Mtr4-RNA exosome interaction <i>in vivo</i>	144
Figure 5. The modeled multiple myeloma amino acid substitution in Rrp4 does not impact Rrp4 protein level or association of the cap subunit with the RNA exosome complex.....	146
Figure 6. The <i>rrp4-M68T</i> mutant cells show specific negative genetic interactions with <i>mtr4</i> mutants that are predicted to impair the Trf4/5-Air1/2-Mtr4 (TRAMP) complex.....	148
Figure 7. The <i>rrp4-M68T mpp6Δ</i> double mutant cells exhibits impaired growth that is exacerbated on drugs that impact RNA processing.....	151
Figure 8. Rrp4 M68T shows decreased association with Mtr4 compared to wild-type Rrp4.....	153
3.8 Chapter III Supplementary Materials.....	154
Supplementary Figure S1. A collection of mutations in RNA exosome subunit genes were identified in newly diagnosed multiple myeloma patients in the CoMMpass study.....	154
Supplementary Figure S2. ConSurf analysis of EXOSC2 and Rrp4 reveals conservation at interface with MTR4/Mtr4.....	156
Supplementary Figure S3. The steady-state level of mature and precursor 5.8S rRNA in <i>rrp4-M68T</i> cells is similar to wild-type, control cells.....	158
Supplementary Figure S4. Synthetical lethality of either <i>rrp4-M68T mtr4-R349E-N352E</i> and <i>rrp4-M68T mtr4-R1030A</i> double mutant cells is rescued by wild-type plasmid.....	160
Supplementary Figure S5. Extended liquid growth curve of <i>rrp4-M68T mpp6Δ</i> and <i>rrp4-M68T rrp47Δ</i> cells.....	161
Supplementary Figure S6. Chr9- NC_000009.12 (130,693,760...130,704,894) schematic with multiple myeloma patient <i>EXOSC2</i> mutations.....	162
Table S1. <i>S. cerevisiae</i> Strains and Plasmids used in this study.....	163
Table S2. DNA Oligonucleotides employed for RT-qPCR.....	165
3.9 References.....	166
Chapter IV. Comparative analysis of disease-linked amino acid substitutions in the RNA exosome modeled in <i>S. cerevisiae</i> reveal functional consequences in translation.....	173
4.1 Abstract.....	174
4.2 Introduction.....	175
4.3 Materials and Methods.....	181
4.4 Results & Discussion.....	184
4.5 Conclusion.....	199
4.6 Chapter IV Figures.....	201

Figure 1. Overview of pathogenic amino acid substitutions in the human cap and core structural subunits of the RNA exosome.	201
Figure 2. <i>S. cerevisiae</i> Rrp4 variants that model EXOSC variants identified in patients show impaired function.	204
Figure 3. RNA-Seq analysis of <i>rrp</i> RNA exosomopathy mutant models reveal distinct transcriptomic changes in the <i>rrp4-G226D</i> , <i>rrp40-W195R</i> and <i>rrp46-L191H</i> cells.	207
Figure 4. UpSet Plots of differentially expressed transcripts in <i>rrp4-G226D</i> , <i>rrp40-W195R</i> and <i>rrp46-L191H</i> cells reveal shared targets involved in metabolism and rRNA processing.	210
Figure 5. UpSet Plots of differentially expressed transcripts in <i>rrp4-G226D</i> , <i>rrp40-W195R</i> and <i>rrp46-L191H</i> cells reveal targets shared differently between dual combinations of the three <i>rrp</i> mutants.	213
Figure 6. UpSet Plots of differentially expressed transcripts in <i>rrp4-G226D</i> , <i>rrp40-W195R</i> and <i>rrp46-L191H</i> cells reveal targets uniquely impacted in each of the three <i>rrp</i> mutants.	216
Figure 7. Heatmaps of <i>rrp</i> mutants reveal broad changes in ribosomal protein gene and CUTs/SUTs expression.	218
4.7 Chapter IV Supplementary Materials	219
Figure S1. MA plots for differential expression analysis results and PCA plots of sample clustering in RNA-Seq experiment.	219
Figure S2. Volcano plots of differentially expressed transcripts identified in <i>rrp40-S87A</i> , <i>rrp45-I15P</i> , <i>rrp46-Q86I</i> and <i>rrp46-L127T</i> samples.	221
Figure S3. Heatmap of ribosomal protein genes with gene names.	224
Table S1. List of RNA exosomopathy mutations and associated pathologies.	225
Table S2. <i>Saccharomyces cerevisiae</i> strains and plasmids	226
Supplemental Documentation S1. Full RNA-Seq Datasets	227
Supplemental Documentation S2. Full lists of differentially expressed genes ($\geq +1.5$ or ≤ -1.5 Fold Change [FC], $p < 0.05$)	227
Supplemental Documentation S3. Full Gene Ontology (GO) Analyses Terms Lists	227
Supplemental Documentation S4. Full Gene Ontology (GO) Analyses on Human homologs	249
4.8 References	262
Chapter V: Discussion & Future Directions.....	269
5.1 Summary of Presented Studies	270
5.2 Conclusions from Presented Studies.....	273
5.3 Future Directions	275
5.4 Closing Remarks.....	281
5.5 Chapter V Figures	283
Figure 1. Proposed model of how RNA exosomopathy amino acid substitutions may result in distinct translational defects that could underlie patient pathology.....	284
5.6 References.....	285

Appendix I. Generating a CRISPR/Cas9 toolkit to introduce RNA exosomopathy-linked missense mutations in <i>S. cerevisiae</i> genes.	290
A1.1 Abstract	291
A1.2 Introduction	292
A1.3 CRISPR/Cas9 Toolkit and Workflow	295
A1.4 Results: Generating and assessing <i>rrp4-G226D</i> and <i>rrp40-W195R</i> mutant cells	306
A1.5 Discussion	309
A1.6 References	317
Appendix II. Experimental Design to assess the impacts of pathogenic amino acid substitutions on RNA exosome integrity	319
A1.1 Introduction	320
A1.2 Experimental Design	322
A1.3 Step I: Generating Inducible Constructs Expressing wild-type Rrp variant with HA tag	324
A1.4 Step II: Verifying expression of generated inducible construct	325
A1.5 Step II:: Timecourse assay to assess exchange rate between the Myc tagged versus the HA tagged Rrp variant.	327
A1.6 Future Directions	331
A1.7 References	336
Appendix III. Amplifying Growth Mindsets With PCR: Implementing growth mindset interventions in introductory biology lab to increase resiliency in undergraduate STEM students.	338
A3.1 Introduction	339
A3.2 Defining “Mindset”	341
A3.3 Rationale and Approach	342
A3.4 Experimental Design	344
A3.5 Results and Limitations of study	350
A3.5 Discussion	353
A3.6 Appendices	355
A3.7 References	367

Chapter I: Introduction

1.1 Post-Transcriptional Regulation of Gene Expression

Though every cell in the human body contains the same DNA “blueprint”, each of us are composed of specialized tissues and organs, distinct in their structures and function. Much of this specialization can be described by differential genetic expression—distinct genes are expressed in different cells, resulting in specific messenger RNAs (mRNAs) that encode the correct proteins for that cell type. While transcriptional control contributes to these distinct gene expression profiles between cells, numerous key regulatory steps after transcription play major roles in achieving differential expression of specific RNAs and affect how those RNAs are made into protein. These steps are categorized as post-transcriptional regulation. Post-transcriptional regulation is essential and involves many conserved factors and non-coding RNAs that play necessary roles in maintaining proper genetic expression. If any of these regulating factors have defects, post transcriptional steps can go awry and lead to serious consequences for human health [1].

As illustrated in Figure 1, post-transcriptional regulation is highly synchronized and consists of a series of steps that occur in both the nucleus and the cytoplasm. While numerous classes of RNA within human cells undergo post-transcriptional regulation, Figure 1 only focuses on post-transcriptional events that occur on mRNAs, from the birth of the nascent transcript to the regulation of decay. Many events depicted are shared with other classes of RNAs, however a few are unique to mRNA regulation, particularly some of the processing events within the nucleus (Figure 1A-C) and translation by the ribosome in the cytoplasm (Figure 1E).

Nuclear post-transcriptional regulatory events (Figure 1A-D):

Most nuclear pre-mRNA processing events occur co-transcriptionally and are extensively coordinated by interactions of the processing factors with the transcriptional machinery. These processing events begin as the mRNA transcript emerges from the polymerase with 5'-end capping (Figure 1A), followed by splicing to remove introns (Figure 1B), and 3'-end cleavage/polyadenylation (Figure 1C). Early in transcription, the 5' end of the nascent mRNA emerges from RNA polymerase II. This 5' end must be protected for the nascent mRNA to mature and survive the transit to translation in the cytoplasm. This protection comes in the form of a 7-methylguanosine cap which is added via 5' end capping (Figure 1A). Eukaryotic mRNA contains this 7-methylguanosine cap, which is added co-transcriptionally to the first 25-30 nucleotides of newly transcribed mRNA by a 5' to 5' triphosphate linkage [2-4]. This unique linkage protects nascent mRNA from nuclear exoribonucleases and also serves as a unique identifier for recruiting other post-transcriptional regulatory factors and translational machinery in the cytoplasm [2].

Once capped, the emerging mRNA undergoes another post-transcriptional event, splicing (Figure 1B). Splicing events remove introns from the nascent transcript via a large macromolecular complex termed the spliceosome [5]. The spliceosome coordinates a series of steps mediated by small nuclear ribonucleoprotein complexes termed U1, U2, U4, U5 and U6 snRNPs [6, 7]. Through a series of protein and RNA rearrangements, the catalytically competent snRNPs perform two trans-esterification reactions that “splice out” the intron from the nascent mRNA [7]. In a coordinated manner, 5' end capping and splicing events occur quickly and provide the foundation for maturation of the nascent mRNA transcript. These two post-transcriptional regulatory steps are of vital importance for the remaining processes that the newly synthesized mRNA must undergo to reach the cytoplasm and be properly translated.

With the 5' end of the newly transcribed mRNA capped, and the introns spliced out, the 3' end must undergo maturation. This maturation occurs through another post-transcriptional regulatory step known as 3' end cleavage and polyadenylation (Figure 1C). The 3' end cleavage and coupled polyadenylation event is critical for mRNA export, stability and even translation in the cytoplasm. Cleavage and polyadenylation facilitate mRNA release from the genomic transcription site and RNA polymerase, ultimately allowing for mRNA export and translation in the cytoplasm. Additionally, poly(A) tails are recognized and bound by poly(A) binding proteins (PABPs) that help protect the newly transcribed mRNA from nuclear decay events and facilitate export from the nucleus [8-10]. Additionally, many other factors are associated with the newly matured mRNA, “packaging” the transcript for export [7, 9]. The packaged mRNA is then targeted for export through nuclear pore complexes (Figure 1D). Proper “packaging” of the mature mRNA strand is required to target the transcript to the pore for efficient export [1, 11]. Failure or stalling of export acts as a signal for RNA decay within the nucleus [12]. Therefore, this key post-transcriptional export step is vital to ensure that only properly processed and packaged mRNAs can access the cytoplasm and interface with ribosomes for translation.

Cytoplasmic post-transcriptional regulatory events (Figure 1E-G):

Once an mRNA reaches the cytoplasm, the primary goal is to translate the information encoded in the mature mRNA into a protein; however, there are mechanisms for both temporal and spatial regulation of translation. To start translation (Figure 1E), the translation initiation complex, eIF4G, recognizes and binds to the mature transcript. This interaction with eIF4G initiates the recruitment of the small and then the large ribosome subunits to the bound mRNA to form a functional ribosome and translate the information in the mRNA into protein. Regulatory mechanisms that can modulate the time and space of translation include mRNA localization to specialized regions for

local translation (Figure 1F) [13] and storage in bodies within the cytoplasm such as P-bodies (Figure 1G) [14]. The fate of the mRNA in the cytoplasm is dynamic with the possibility to move between storage, localization, and translation.

1.2 RNA Turnover and Decay

The final post-transcriptional regulatory event depicted in Figure 1, RNA turnover (Figure 1H), regulates multiple aspects that dictate steady-state transcript levels. How mRNA turnover can occur if mature mRNA is not properly packaged in the nucleus, resulting in transcript decay, has already been mentioned; however, mRNA turnover can also occur if any of the other post-transcriptional steps discussed thus far go awry—if the transcript is improperly spliced, capped, or polyadenylated, the nascent mRNA will be targeted for decay within the nucleus. If these previous post-transcriptional events occur correctly, and the mRNA is properly processed, packaged and exported, turnover can be tuned as a means of regulating the amount of protein produced. After multiple rounds of translation, the poly(A) tail becomes shorter and the mRNA can be de-capped by cytoplasmic enzymes. This decapping signals to the cell that the mRNA needs to be destroyed by cytoplasmic RNA decay machinery. These regulatory decay steps can therefore modulate the steady-state level of the message and control the amount of protein generated via translation. Given the overarching impact RNA turnover has in regulating mRNA, it is no surprise cells have a variety of evolutionary conserved decay pathways and mechanisms to control this process [15].

In addition to mRNA turnover and decay, surveillance of noncoding RNA species through regulated decay is essential for cellular health and function. Noncoding RNA surveillance occurs in both the cytoplasm and the nucleus by a collection of exo- and endo- ribonucleases [16]. As exoribonucleases require sing-stranded accessible ends, the accessibility of noncoding RNAs is a

major determinant of decay. Several noncoding RNAs are structured and bound to chaperone proteins to protect from ribonuclease targeting; thus, a failure to bind chaperones or aberrant folding of a noncoding RNA would be a signal for decay. Ribonucleases can target noncoding RNAs through interacting cofactor that recognize and/or assist in generating accessible 3' or 5' ends [16]. Endoribonucleases can also aid exoribonuclease mediated decay by cleaving structured ncRNA and generating accessible ends for targeted degradation [16]. Additionally, some interacting cofactors recognize and bind to sequences of the target or have helicase activity to unwind the structured noncoding RNA [16].

RN turnover and decay of both coding and noncoding species are regulated through conserved, overlapping pathways [16]. One such mechanism that contributes to many RNA decay and turnover events throughout the cell is the RNA exosome complex. The RNA exosome is a multisubunit 3'-5' exonuclease that both processes and destroys numerous classes of noncoding RNAs and mRNAs [15, 17]. The complex consists of ten highly conserved and essential subunits, nine non-catalytic structural components and one catalytic exo/endonuclease [18]. Through interactions with cofactor proteins or complexes, the RNA exosome targets transcripts in both the cytoplasm and nucleus [19]. Within the nucleus, interacting cofactors aid the exosome in monitoring "RNA homeostasis", targeting aberrant mRNA and noncoding transcripts [12]. Within the cytoplasm, interacting cofactors target transcripts that are no longer useful or are inefficiently translated for decay by the RNA exosome [7, 12, 20]. In addition to the RNA exosome, eukaryotic cells have many pathways that can target RNA for decay in a 5'-3' direction. The major 5'-3' exoribnuclease involved in RNA decay belong to the XRN family [16, 21]. There are both cytoplasmic and nuclear XRN exoribonucleases that monitor many classes of RNA. Similar to the RNA exosome, many of these decay events are regulated through interactions with cofactor

proteins that aid in targeting the exoribonucleases [21]. In many instances, XRN exoribonucleases and the RNA exosome work in tandem to decay cytoplasmic mRNAs [20, 22, 23]. This demonstrates the redundancy in decay pathways that exist within eukaryotic life. However, this redundancy should not be attributed to “evolutionary oversight”, but rather to the vital role RNA decay has in regulating RNA throughout its lifespan and ours.

The essential role for decay and turnover in establishing proper genetic expression is underscored by the links to defects in RNA decay machinery and human disease. Multiple disease-linked mutations have been identified in components of the RNA exosome [18, 24]. Mutations in the gene encoding the catalytic component of the RNA exosome have been linked to multiple myeloma [24], however surprisingly disease mutations in genes encoding the non-catalytic, structural components have also been identified [25-31]. Though these mutations are within genes that encode a ubiquitous, conserved molecular machine, the diseases manifest in specific tissues. There is little information about any tissue-specific functions of the RNA exosome and given its critical roles in both RNA decay and RNA processing of nearly every class of RNA, RNA exosome function is unlikely to substantially differ between different human tissues. Thus, the question arises how defects in this key, post-transcriptional regulatory machine can result in tissue specific defects.

1.3 The RNA Exosome: An Essential Molecular Machine

Differential genetic expression allows for the cellular diversity and complex organization of distinct tissues within our bodies. Key to this cell-specific genetic expression are post-transcriptional events that regulate RNA levels [1]. The RNA exosome is a critical post-transcriptional regulator, required for 3' to 5' processing and degradation of a vast number of RNAs in both the nucleus and cytoplasm [19, 32-34]. First identified in a budding yeast genetic

screen for ribosomal RNA processing (*rrp*) mutants [35, 36], the RNA exosome has since been shown to contribute to the processing and/or degradation of nearly every class of RNA [34, 37-39] including mRNAs, small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), transfer RNAs (tRNAs), and transcripts that result from pervasive transcription such as cryptic unstable transcripts (CUTs) [24, 32, 40-42]. The diversity of RNA targets implicates the conserved complex in several biological processes and illustrates the influence the RNA exosome has on cellular function.

Structural studies of the RNA exosome reveal a conserved 9-subunit core common to eukarya and archaea (Figure 2) [43-54]. The common core of the RNA exosome consists of an upper cap comprised of three subunits containing RNA binding S1 and K homology (KH) domains atop a lower hexameric ring of RNase PH subunits. Within eukarya, the nine structural subunit core associates with a 3'-5' exo/endoribonuclease [DIS3 (human); Rrp44 (budding yeast)] [35, 45] as the RNase PH subunits have lost phosphorolytic active sites and are catalytically inert [43, 48] (Figure 2A). The eukaryotic RNA exosome cap is composed of three evolutionarily conserved S1/KH cap subunits, EXOSC1/2/3 in humans (Figure 2B) or Csl4/Rrp4/Rrp40 in budding yeast (Figure 2C). The lower ring of the eukaryotic RNA exosome consists of six conserved PH-like subunits, EXOSC4/7/8/9/5/6 in humans (Figure 2B) or Rrp41/Rrp42/Rrp43/Rrp45/Rrp46/Mtr3 in budding yeast (Figure 2C). The archaeal RNA exosome consists of a S1/KH cap containing a trimer of a singular EXOSC2/Rrp4 homolog and a catalytically active RNase PH ring composed of three dimers of EXOSC4/Rrp41 and EXOSC7/Rrp42 homologs [48, 52] (Figure 2D). As illustrated in Figure 2, the 9-subunit common core of the RNA exosome forms a barrel like structure, with a central pore in the S1/KH cap through which RNA target substrates can be threaded to access the catalytic sites, either the associated exo/endoribonuclease in eukarya or the

RNAse PH-ring in archaea [32, 48]. While the mechanisms of RNA degradation/processing have diverged in archaea and eukarya, these structural studies suggest evolutionary conservation within the organization of the RNA exosome across all life. Moreover, functional studies have revealed that the RNA exosome is essential in all eukaryotic systems studied thus far [35, 39, 48, 55, 56], further highlighting the pivotal role this molecular machine plays in modulating levels of gene expression.

1.4 Function of the Eukaryotic RNA Exosome

This pivotal role in modulating levels of gene expression is made more apparent when cataloging the function of the RNA exosome in eukarya. As mentioned, the RNA exosome contributes to the processing and/or degradation of nearly every class of RNA in both compartments of the cell [32, 34, 37-40, 42, 57, 58]. A key role of the nuclear eukaryotic RNA exosome is processing ribosomal RNA (rRNA), particularly the precursor 5.8S rRNA [35]. The 5.8S rRNA is one of 3 rRNAs, along with several ribosomal proteins, that make up the large 60S subunit of the eukaryotic ribosome. Within the nucleolus, the 5.8S rRNA is transcribed as a single, large precursor rRNA containing the other rRNAs that will make up the eukaryotic ribosome. This long precursor rRNA is cleaved and processed by various endo- and exonucleolytic reactions [59, 60]. The eukaryotic RNA exosome trims the 3' end of the cleaved 5.8S precursor, generating the mature 5.8S species that makes up the large ribosomal subunit [36]. The level of mature 5.8S rRNA will dictate the number of functional ribosomes that can exist in the cell. Thus, RNA exosome function can greatly impact translation.

Another key role of the RNA exosome is the degradation of RNAs that arise from cryptic or pervasive transcription [32]. These are highly unstable species that usually are only detectable when RNA surveillance is compromised [32]. The exact function of many of these cryptic,

unstable transcripts remains unknown, however several studies have linked these RNA species to regulating gene expression [61-64]. Thus tight regulation of these level of these cryptic transcripts by RNA exosome likely helps maintain transcriptomic homeostasis in the eukaryotic cell. The RNA exosome also plays roles in mRNA quality control as described previously. In the nucleus, the RNA exosome degrades pre-mRNAs with retained introns, pre-mRNAs that are improperly packaged for export, and those that failed to transcriptionally terminate at the proper polyadenylation site [32]. In the cytoplasm the RNA exosome acts in the mRNA decay pathways nonsense mediated decay (NMD), nonsense decay (NSD) and no-go decay (NGD) [33].

In addition to these well-established roles, RNA exosome function has recently been linked to DNA damage repair pathways, resolving R loops, and telomerase RNA quality control [32, 65]. Recent studies also suggest the nuclear RNA exosome is essential for maintaining progenitor cell function and preventing premature stem cell differentiation [66]. Overall, the RNA exosome truly is a molecular machine central to proper cellular function. The set of targets the RNA exosome degrades or processes broadly impacts several biological processes and thus expands the complex's influence in eukaryotic cellular health.

1.5 Interacting Cofactors Expand the Function of the RNA Exosome

The specificity of the RNA exosome for the broad set of target transcripts is conferred by several interacting cofactors [65]. Cofactor interactions with the eukaryotic RNA exosome aid the complex in target recognition, RNA unwinding, degradation, and catalysis in both the nucleus and the cytoplasm [65]. Many RNA exosome cofactors were first identified and characterized in budding yeast. The nuclear RNA exosome associates with an additional 3'-5' exonuclease Rrp6 that associates with the complex and aids in nuclear RNA targeting and processing [46, 67]. Rrp6

has an obligate binding partner, Rrp47, which stabilizes the interaction of the Rrp6 exonuclease with the RNA exosome complex [65, 68]. Additionally, a small, intrinsically disordered protein, Mpp6, interacts with the RNA exosome through contacts with cap subunits and aids the complex in nuclear RNA surveillance [69, 70]. The essential 3' to 5' DExH box RNA helicase Mtr4 has a central role in nuclear RNA turnover by the RNA exosome [71-73]. The Mtr4 helicase assists in RNA substrate unwinding and plays a critical role in RNA exosome processing of the 5.8S rRNA precursor (7S rRNA) [73, 74]. Orthologs of these yeast nuclear cofactors have since been identified in the mammalian system [EXOSC10 (human Rrp6), C1D (human Rrp47); MPH6 (human Mpp6); MTR4/MTREX (human Mtr4)] [65]. Structural studies of the budding yeast and mammalian RNA exosome reveal that Rrp6/EXOSC10, Rrp47/C1D and Mpp6/MPH6 interact with the complex through conserved interfaces that form composite sites for interactions with other cofactors such as Mtr4/MTR4/MTREX [75-79]. Mtr4/MTR4/MTREX also acts as part of larger complexes that aid the RNA exosome in nuclear RNA quality control, including the budding yeast TRAMP (Trf4/5-Air1/2-Mtr4 Polyadenylation) complex and the mammalian NEXT (Nuclear Exosome Targeting) complex [21, 32, 80-84]. Therefore, these conserved cofactor interactions with the RNA exosome not only impact the function of the complex but expand the molecular machine's interactome and influence within the cell.

1.6 RNA Exosome and Human Disease: RNA Exosomopathies

Many different human disease mutations have been identified in genes encoding RNA exosome subunits [24]. Recurrent somatic mutations in *DIS3*, which encodes the catalytic component of the human RNA exosome [85], have been linked to multiple myeloma [86, 87]. Multiple myeloma, which is the second most common hematologic malignancy, is an incurable

cancer of the long-lived antibody-secreting plasma cells of the bone marrow [88, 89]. While multiple myeloma is a genetically heterogeneous disease that is confounded by chromosomal translocations and copy number variants, *DIS3* mutations are the fourth most common single nucleotide variation identified in multiple myeloma [86, 87]. Multiple myeloma-associated *DIS3* mutations disrupt proper RNA degradation and processing in both mammalian cells and budding yeast mutant cells [90-92]. However additional mechanistic studies are required to understand how mutations in *DIS3*, and the function of the RNA exosome, could contribute to pathogenesis in multiple myeloma.

Recent studies have also linked missense mutations in *EXOSC* genes encoding the structural subunits of the RNA exosome to various human pathologies, comprising a growing family of diseases termed “RNA exosomopathies” [24, 29-31, 93-98]. Intriguingly, these RNA exosomopathy mutations cause a range of clinical manifestations that do not appear to correlate with the type of structural gene that is mutated. Mutations in the cap subunit genes *EXOSC3* and *EXOSC1*, and the core subunit gene *EXOSC8*, have been linked to forms of pontocerebellar hypoplasia type 1, a neurological disorder characterized by atrophy of the pons and cerebellum [31, 33, 93, 94, 96-98]. Mutations in the core subunit genes *EXOSC5* and *EXOSC9* have been linked to similar neurological defects including cerebellar degeneration, neuronopathy and neurodevelopmental delays [29, 30, 99]. In contrast to the other exosomopathy mutations described thus far, missense mutations in the cap subunit gene *EXOSC2* have been linked to a novel, complex disorder characterized by retinitis pigmentosa, progressive hearing loss, premature aging, short stature, mild intellectual disability and distinctive gestalt [95], later named SHRF (Short stature, Hearing loss, Retinitis pigmentosa and distinctive Facies) (OMIM #617763) [100].

While diverse in clinical manifestations, all of these identified RNA exosomopathy mutations result in single amino acid substitutions that occur in highly conserved domains of the RNA exosome subunits (Figure 3). Explaining the tissue-specific nature of these different RNA exosomopathy diseases is challenging if the amino acid substitutions generally affect the molecular function(s) of this complex. Rather, the diverse clinical phenotypes linked to these different RNA exosomopathy mutations suggest that each amino acid substitution has specific molecular consequences that differentially impact the function of the complex. The link between *DIS3* single nucleotide variants and multiple myeloma further adds evidence that human disease mutations in RNA exosome genes can result in differential defects *in vivo*. Taken together, these clinical studies and the ongoing identification of disease-linked RNA exosome mutations raises the question as to how single amino acid substitutions in highly conserved subunits of an essential complex can differentially affect the function of the RNA exosome and consequently underlie distinct, tissue specific clinical manifestations.

1.7 RNA Exosomopathy-linked Amino Acid Substitutions and Differential Effects on RNA exosome Function *in vivo*

The growing number of RNA exosomopathies with distinct clinical presentations suggests that RNA exosome mutations have differential *in vivo* effects. These disease-linked amino acid substitutions, as modeled in Figure 4, could differentially affect (I) the integrity of the complex, (II) key interactions, and/or (III) directly disrupt processing/degradation of RNAs by this molecular machine. Recent work reports changes in overall levels of the RNA exosome complex in *EXOSC9*-linked RNA exosomopathy patient fibroblasts [30], thus suggesting that distinct amino acid substitutions could deplete the complex; however, this observation from fibroblasts is difficult

to reconcile with patients that show primarily neurological deficits. Other work shows that an EXOSC3 disease-linked amino acid substitution modeled in *S. cerevisiae* decreases the affinity for a specific RNA exosome cofactor Mpp6 [76], suggesting that an RNA exosomopathy amino acid change could affect key interactions of the RNA exosome necessary for proper function. There could also be a dynamic relationship between cofactor interaction loss and decreased complex integrity resulting from some of these RNA exosomopathy amino acid substitutions. How the structural subunits of the RNA exosome come together to assemble the complex with precise stoichiometry remains unknown. Interacting cofactors could act as chaperones to mediate the assembly of the RNA exosome core similar to events that occur with other conserved, multi-subunit complexes such as the proteasome [101, 102]. Therefore, amino acid changes in these structural subunits that destabilize an interaction could also impact the assembly of the complex. Likewise, a complex that is thermodynamically unstable due to amino acid changes in structural components would disrupt cofactor associations as composite sites on the RNA exosome that mediate interactions may not be formed.

Beyond effects on complex integrity, different amino acid substitutions in structural core subunits could disrupt the central pore through which RNA substrate is threaded to DIS3, the catalytic exo/endoribonuclease that degrades/processes target transcripts (Figure 2) [32, 48]. While no structures of the RNA exosome containing subunits with pathogenic amino acid substitutions have been solved, recent structural studies have visualized the path nucleic acid substrate takes through the human RNA exosome at the atomic level [44, 79]. These studies revealed that single-stranded RNA interacts directly with EXOSC2 residues facing the central pore of the S1/KH cap and with conserved residues of EXOSC4 and EXOSC9 at the base of the central

channel within the PH-like core ring [44, 79]. Therefore, amino acid changes within structural subunits could biochemically disrupt the path RNA substrate takes in the human RNA exosome.

Ultimately, disruption of cofactor interactions, complex integrity or the central channel of the complex will compromise processing and/or degradation of RNA exosome targets. Any alteration in the RNA exosome function resulting from these amino acid substitutions could have a profound impact in certain tissues if key classes or specific RNAs are misprocessed or if defective RNA accumulates. Thus, to better understand the disease mechanism of these exosomopathies, understanding how these different pathogenic amino acid substitutions affect the function of the RNA exosome *in vivo* is essential.

1.8 Doctoral Research Aims

In the following text, I present my doctoral work addressing how these different pathogenic amino acid substitutions affect the function of the RNA exosome by characterizing and comparatively analyzing these RNA exosomopathy-linked changes *in vivo*. Using the *Saccharomyces cerevisiae* genetic model system, I modeled patient RNA exosomopathy mutations in orthologous genes and assessed the *in vivo* consequences resulting the resulting amino acid change. I first present a systematic approach of characterizing a single RNA exosomopathy-linked amino acid change identified in EXOSC2 (Chapter II). I also present work using the same approach, characterizing a different human disease mutation found in *EXOSC2* that resulted in distinct *in vivo* consequences from those caused by the RNA exosomopathy pathogenic substitution (Chapter III). Lastly, I present a comparative transcriptomics approach to assess multiple *S. cerevisiae* RNA exosomopathy mutant models to understand if there are differential impacts on the complex's function from each of these disease-linked mutations. With this

comparative approach, I integrate my analysis with other RNA exosome mutants such as the *EXOSC3* and *EXOSC5* models that were studied previously [29, 103, 104]. The approaches I present here using *S. cerevisiae* models allow for a comprehensive analysis of the *in vivo* molecular underpinnings of each RNA exosomopathy mutation given the genetic tools available and the extensive historical knowledgebase of studying the RNA exosome in budding yeast [35, 40, 57, 58, 97, 105-107]. Furthermore, the *S. cerevisiae* system is a tractable model for mentorship opportunities, which is evident by the numerous undergraduate mentee authors who contributed to each chapter. This project further establishes the use of *S. cerevisiae* to model human disease-associated mutations in the RNA exosome and will act as an example as to how systematically, and comparatively, survey the *in vivo* consequences of the growing list of identified RNA exosomopathy mutations.

1.9 Chapter I Figures

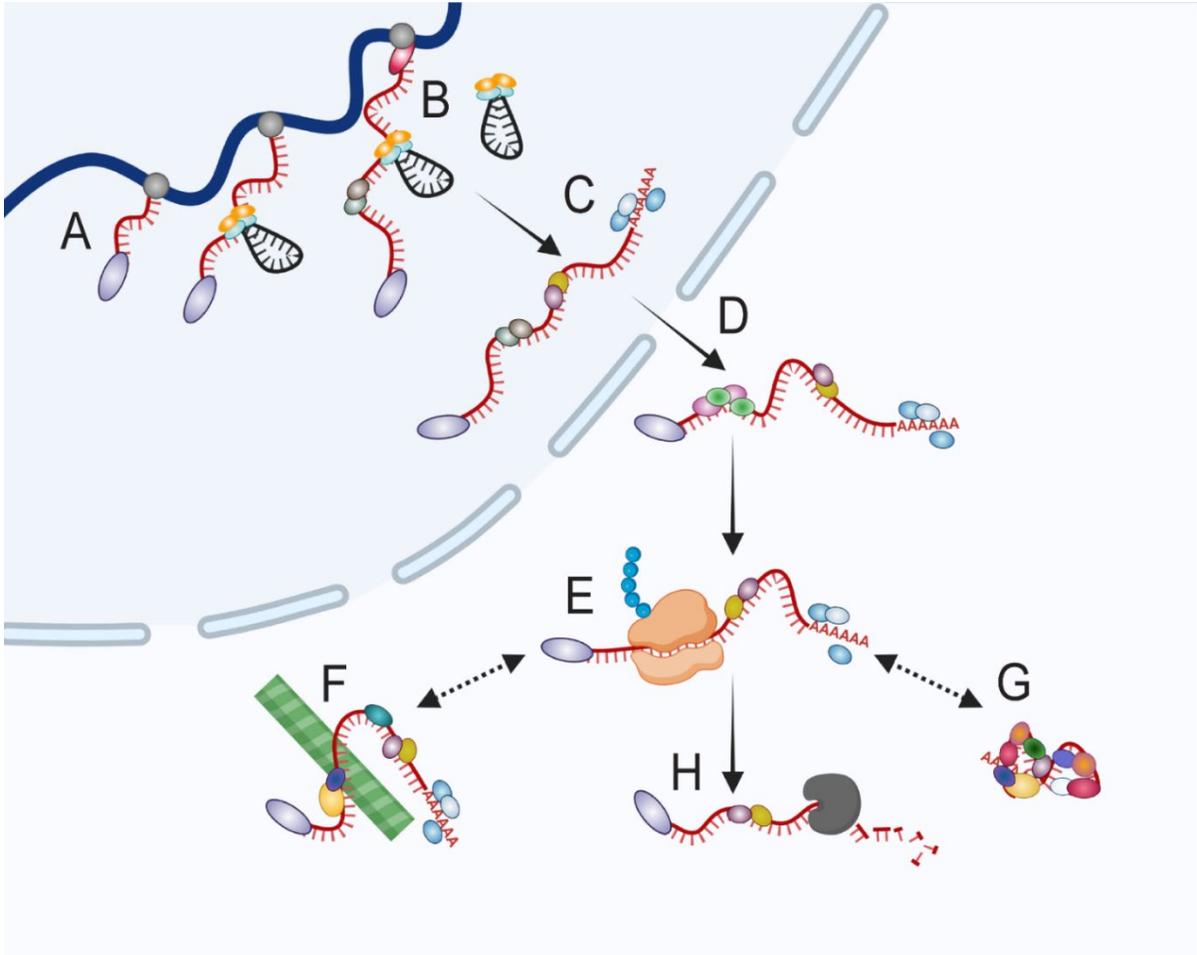


Figure 1. Schematic representation of mRNA post-transcriptional regulation and gene expression.

(A) As nascent mRNA emerges from RNA polymerase II, capping enzymes add a 7-methylguanosine cap to the emerging 5' end of the transcript via a 5' to 5' triphosphate linkage.

(B) As transcription continues, splicing machinery binds to the transcript and introns are co-transcriptionally removed. (C) At the 3' end of the transcript, 3'-end cleavage and polyadenylation are coupled with the release of the mRNA from RNA polymerase II. (D) The mature transcript is

then properly processed and packaged for export to the cytoplasm. Once in the cytoplasm, the

mRNA can be (E) translated to produce the encoded protein; (F) localized to specific locations within the cells and/or (G) stored in cytoplasmic bodies such as P-bodies. These processes are dynamic, meaning a mature mRNA can be localized and then translated or move in and out of storage bodies depending on the need for the encoded message and/or the requirements of the cellular environment. Ultimately, the transcript undergoes (H) mRNA turnover and decay. Figure modeled after *Sterrett and Corbett* 2019 eLS chapter “Post-Transcriptional Regulation of Gene expression and Human Genetic Disease” [108].

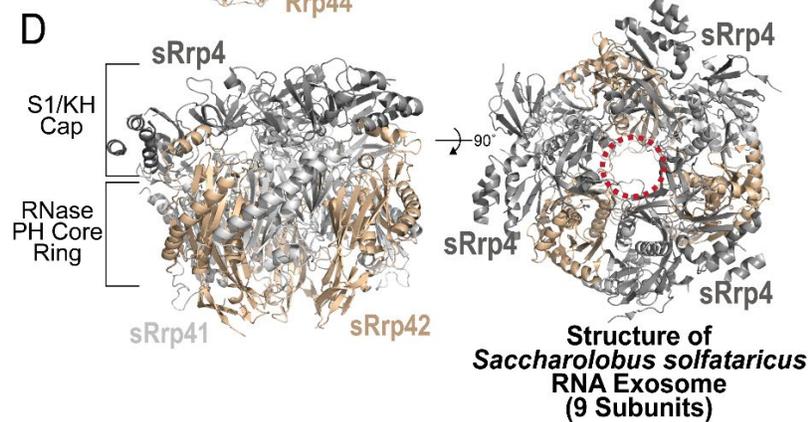
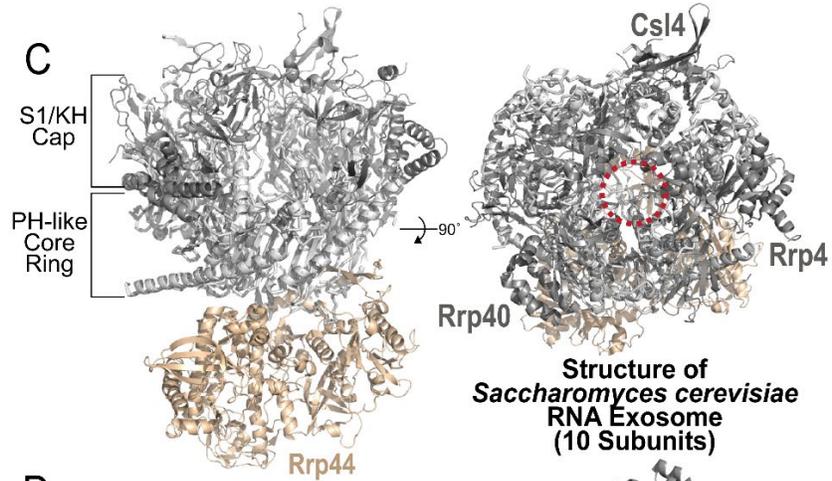
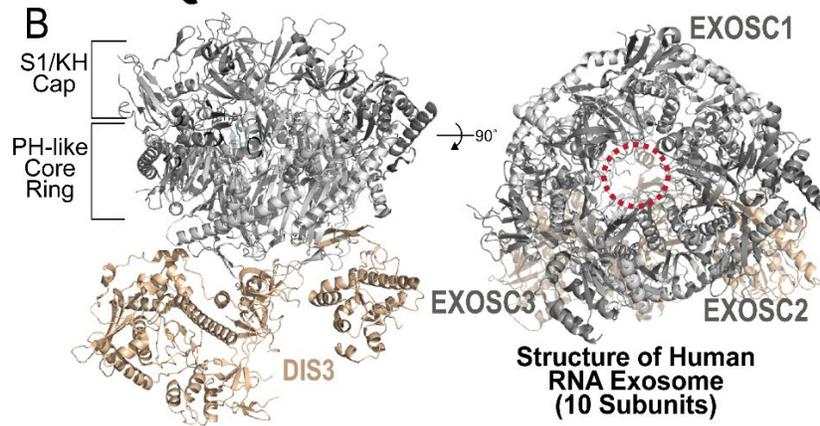
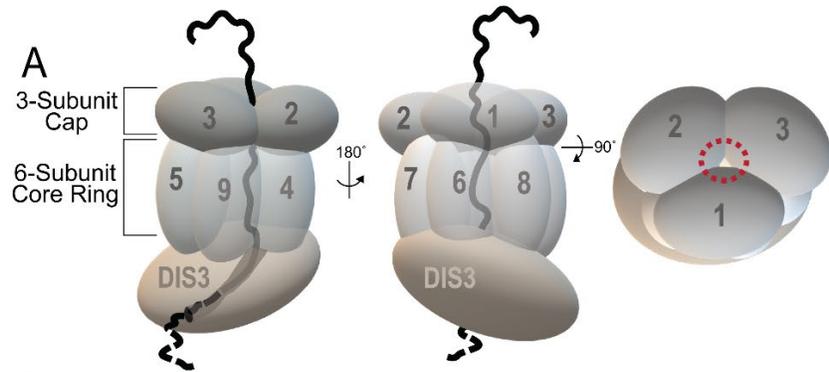


Figure 2. Conservation of the core RNA exosome structure in eukarya and archaea.

(A) Cartoon schematic of the eukaryotic RNA exosome. The RNA exosome is an evolutionary conserved ribonuclease complex composed of nine structural subunits (EXOSC1-9) one catalytic subunit (DIS3) that form a “cap” and “core” ring-like structure. A pore within the cap (outlined by dashed red circle) is the opening of a central channel through which RNA substrate is threaded to reach the active sites in the catalytic subunit. (B) Structure of the human RNA exosome (PDB 6D6Q) [79]. The 3-subunit cap at the top of the complex is composed of EXOSC1, EXOSC2 and EXOSC3. The 6-subunit core is composed of EXOSC4, EXOSC5, EXOSC6, EXOSC7, EXOSC8, and EXOSC9. The DIS3 catalytic subunit is located at the bottom. The cap and core subunits contain conserved protein domains with EXOSC1/2/3 containing putative RNA binding domains S1 and KH and EXOSC4/5/6/7/8/9 containing RNase PH-like domains. The conserved pore through which RNA substrate is threaded is outlined by a dashed red circle. (C) Structure of the *Saccharomyces cerevisiae* RNA exosome (PDB 6FSZ) [109]. The 3-subunit cap at the top of the complex is composed of Csl4, Rrp4 and Rrp40. The 6-subunit core is composed of Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and Mtr3. The Rrp44 catalytic subunit is located at the bottom. The cap and core subunits contain conserved protein domains with Csl4/Rrp4/Rrp40 containing putative RNA binding domains S1 and KH and Rrp41/Rrp42/Rrp43/Rrp45/Rrp46/Mtr3 containing RNase PH-like domains. The conserved pore through which RNA substrate is threaded is outlined by a dashed red circle. (D) Structure of the archaeal *Saccharolobus solfataricus* RNA exosome (PDB 4BA1) [52]. The 3-subunit cap at the top of the complex is composed of a trimer of sRrp4, a homolog of the human EXOSC2 and yeast Rrp4 cap subunit. The 6-subunit core ring is composed of three dimers of sRrp41 and sRrp42, homologs of the human core subunits EXOSC4/7 and the yeast core subunits Rrp41/Rrp42. Similar to EXOSC2/Rrp4, sRrp4 contains S1/KH RNA binding domains.

The core sRrp41 and sRrp42 subunits contain catalytically competent RNase PH domains. Similar to the eukaryotic RNA exosomes, the archaeal complex has a central pore (outlined by dashed red circle) through which RNA substrate is threaded to reach the active sites of sRrp41/sRrp42.

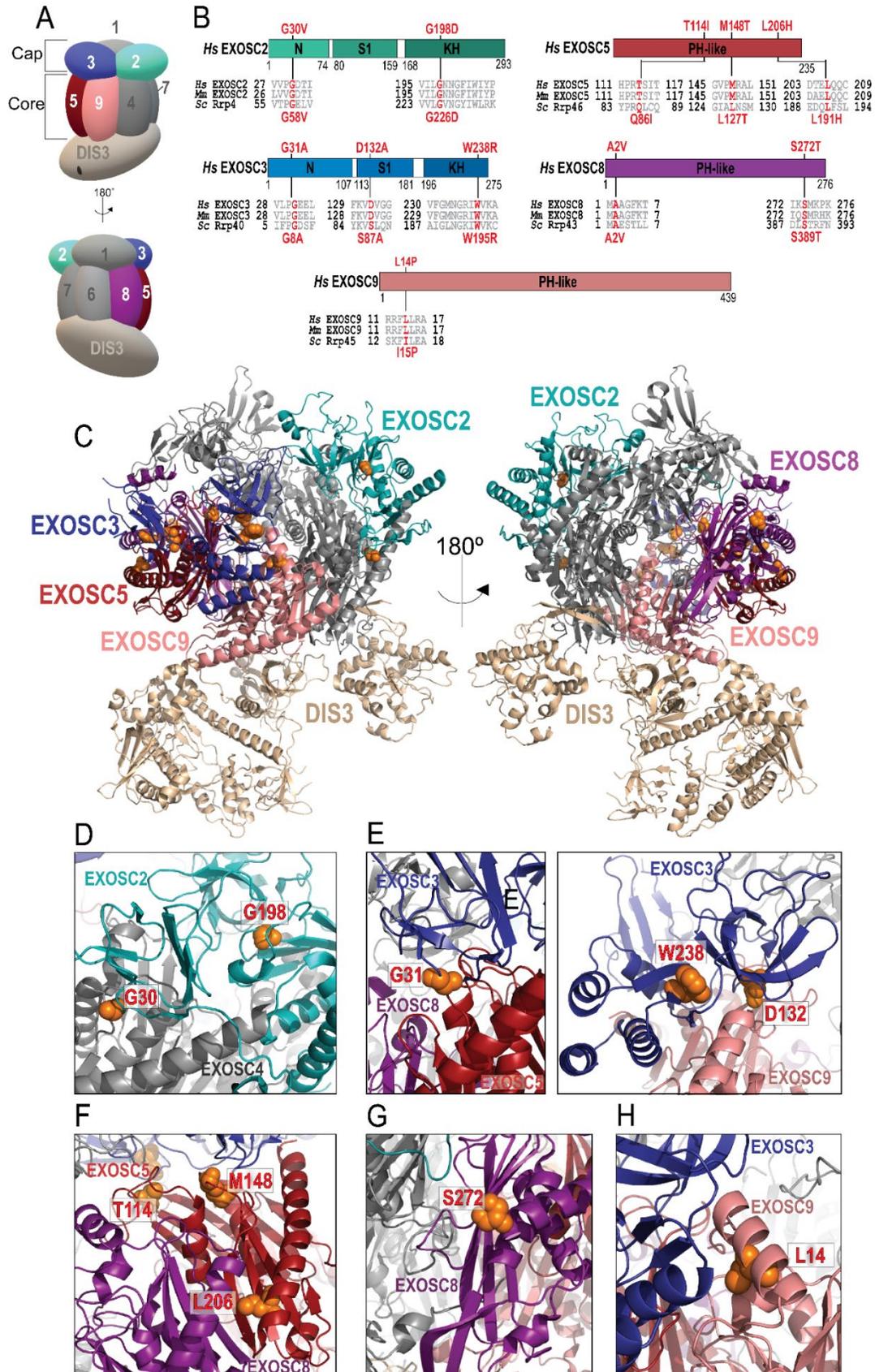


Figure 3. Overview of pathogenic RNA exosomopathy amino acid substitutions in the human RNA exosome.

(A) Cartoon of the human RNA exosome. Missense mutations in the genes encoding the colored subunits have been identified in RNA exosomopathy patients. Pathogenic amino acid substitutions in the cap subunit EXOSC2 (teal) cause a novel, complex disorder SHRF (Short stature, Hearing loss, Retinitis pigmentosa and distinctive Facies) (OMIM #617763) [95, 100]. Pathogenic amino acid substitutions in the cap subunit EXOSC3 (blue) and core subunit EXOSC8 (purple) are linked to forms of pontocerebellar hypoplasia (PCH1b and 1c, respectively) [33, 93, 94, 96-98], while pathogenic substitutions in core subunits EXOSC5 (red) and EXOSC9 (pink) have been linked to a similar neurological defects including cerebellar degeneration, neuropathy and neurodevelopmental delays [29, 30, 99]. (B) Domain structures are shown for EXOSC2, EXOSC3, EXOSC5, EXOSC8 and EXOSC9. The cap subunits EXOSC2 and EXOSC3 are composed of three different domains: an N-terminal domain, an S1 putative RNA binding domain, and a C-terminal putative RNA binding KH (K homology) domain. The “GxNG” motif identified in the KH domain of both cap subunits is boxed in green. The core subunits EXOSC5, EXOSC8 and EXOSC9 are composed of a single domain that resembles a RNase PH domain. The position of the disease-linked amino acid substitutions in the human subunits are depicted above the domain structures in red. Sequence alignments of the EXOSC orthologs from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*) and *S. cerevisiae* (*Sc*) below the domain structures show the highly conserved residues altered in disease in red and the conserved sequences flanking these residues in gray. The amino acid substitutions in *S. cerevisiae* Rrp subunits generated for the doctoral work presented here are shown below the structures in red. (C-H) The human RNA exosome mapped with the conserved residues substituted in RNA exosomopathy patients that are visualized in the structure

(PDB 6D6Q) [79]. The conserved residues substituted within RNA exosomopathy patients are depicted as orange spheres. (D) EXOSC Gly30 (G30) and Gly198 (G198) are altered in SHRF patients. EXOSC2 G30 is located between EXOSC2 and EXOSC4. EXOSC2 G198 is at the end of a β -strand surrounded by four β sheets, suggesting EXOSC2 G198 plays a structural role within the subunit. (E) EXOSC3 Gly31 (G31), EXOSC3 Trp238 (W238) and EXOSC3 Asp132 (D132) are substituted in PCH1b patients and may impact intersubunit interactions. EXOSC3 G31 is substituted to an Alanine in PCH1b patients. EXOSC3 31 located at the surface of EXOSC3 near EXOSC5. EXOSC3 W238 and EXOSC3 D132 are at the interface of EXOSC3 and EXOSC9. (F) Residues Thr114 (T114), Met148 (M148) and Leu206 (L206) in EXOSC5 are substituted in RNA exosomopathy patients with neurological defects. EXOSC5 T114 and EXOSC5 M148 are in proximity with neighboring subunit EXOSC3. EXOSC5 L206 is buried within a hydrophobic pocket of the core subunit. (G) EXOSC8 Ser272 (S272) is substituted in PCH1c patients. EXOSC8 272 is located near the bottom of the central channel of the human RNA exosome. A substitution of EXOSC8 272 may biochemically impact the central channel that threads RNA substrate. (H) EXOSC9 Leu14 (L14) is substituted in RNA exosomopathy patients who present with neurological defects. EXOSC9 L14 is located in the first alpha helix of EXOSC9 and could stabilize structural interactions within the subunit.

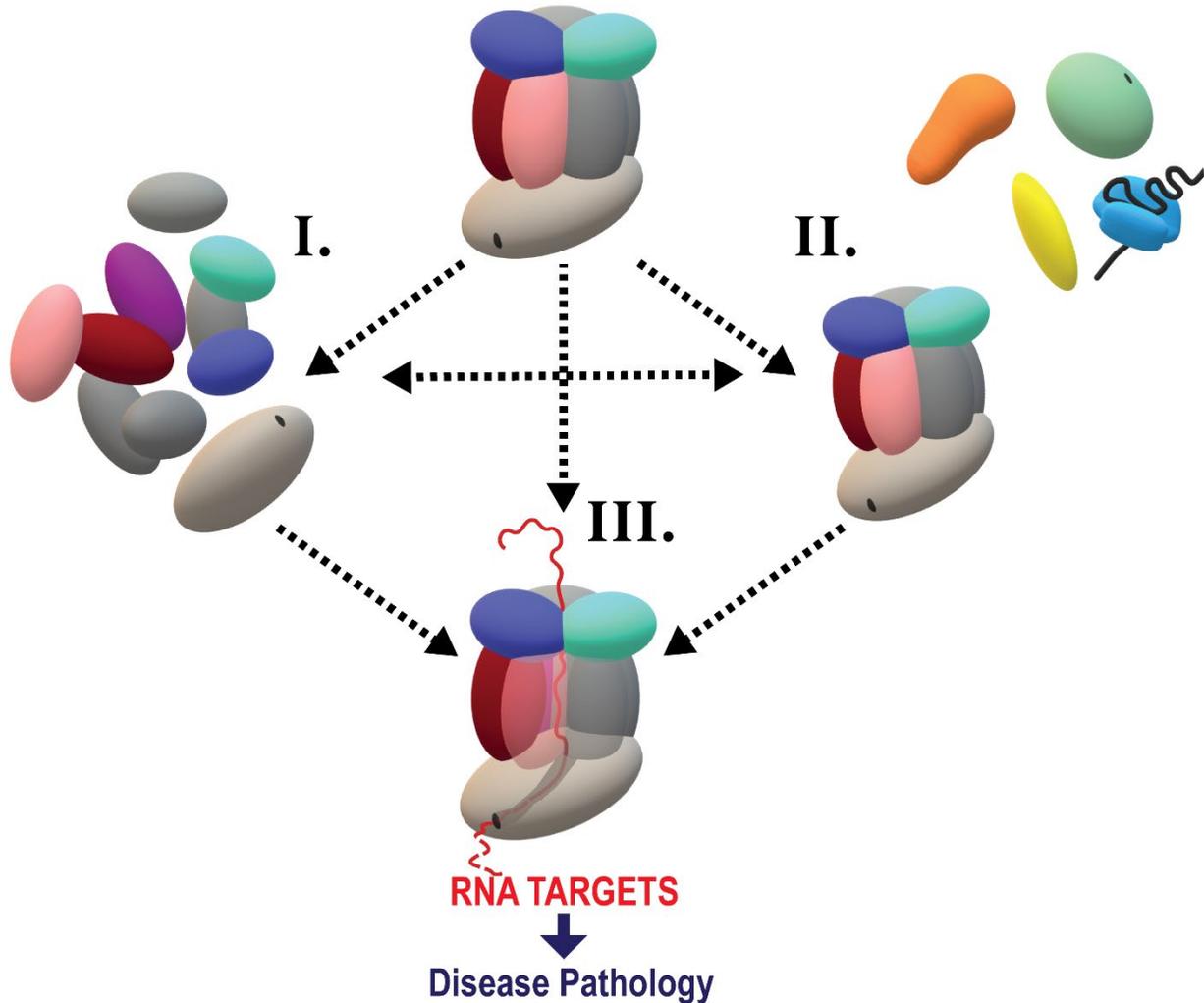


Figure 4. Proposed model of how RNA exosomopathy amino acid substitutions may result in distinct disease pathologies.

Disease-linked amino acid substitutions may differentially affect (I) complex integrity, (II) cofactor interactions, and/or (III) directly disrupt the processing or degradation of transcripts via impacting the central channel. Ultimately these changes will disrupt the RNA targeting and function of the molecular machine, leading to distinct *in vivo* molecular consequences resulting from each different RNA exosomopathy mutation. These differential consequences could underlie the differing disease pathologies observed within RNA exosomopathy patients.

1.10 References

1. Corbett, A.H., *Post-transcriptional regulation of gene expression and human disease*. Curr Opin Cell Biol, 2018. **52**: p. 96-104.
2. Ramanathan, A., G.B. Robb, and S.-H. Chan, *mRNA capping: biological functions and applications*. Nucleic Acids Research, 2016. **44**(16): p. 7511-7526.
3. Moteki, S. and D. Price, *Functional coupling of capping and transcription of mRNA*. Molecular Cell, 2002. **10**(3): p. 599-609.
4. Shatkin, A.J. and J.L. Manley, *The ends of the affair: Capping and polyadenylation*. Nature Structural Biology, 2000. **7**(10): p. 838-842.
5. Shi, Y., *Mechanistic insights into precursor messenger RNA splicing by the spliceosome*. Nature Reviews Molecular Cell Biology, 2017. **18**(11): p. 655-670.
6. Papasaikas, P. and J. Valcarcel, *The Spliceosome: The Ultimate RNA Chaperone and Sculptor*. Trends in Biochemical Sciences, 2016. **41**(1): p. 33-45.
7. Moore, M.J. and N.J. Proudfoot, *Pre-mRNA processing reaches back to transcription and ahead to translation*. Cell, 2009. **136**(4): p. 688-700.
8. Cowling, V.H., *Regulation of mRNA cap methylation*. Biochem J, 2009. **425**(2): p. 295-302.
9. Valdés-Flores, J., et al., *Life and Death of mRNA Molecules in Entamoeba histolytica*. Frontiers in cellular and infection microbiology, 2018. **8**: p. 199-199.
10. Neve, J., et al., *Cleavage and polyadenylation: Ending the message expands gene regulation*. RNA Biol, 2017. **14**(7): p. 865-890.
11. Gerstberger, S., M. Hafner, and T. Tuschl, *A census of human RNA-binding proteins*. Nat Rev Genet, 2014. **15**(12): p. 829-45.
12. Schmid, M. and T.H. Jensen, *Controlling nuclear RNA levels*. Nat Rev Genet, 2018. **19**(8): p. 518-529.
13. Donlin-Asp, P.G., W. Rossoll, and G.J. Bassell, *Spatially and temporally regulating translation via mRNA-binding proteins in cellular and neuronal function*. FEBS Lett, 2017. **591**(11): p. 1508-1525.
14. Standart, N. and D. Weil, *P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage*. Trends Genet, 2018. **34**(8): p. 612-626.

15. Garneau, N.L., J. Wilusz, and C.J. Wilusz, *The highways and byways of mRNA decay*. Nat Rev Mol Cell Biol, 2007. **8**(2): p. 113-26.
16. Belair, C., S. Sim, and S.L. Wolin, *Noncoding RNA Surveillance: The Ends Justify the Means*. Chemical Reviews, 2018. **118**(8): p. 4422-4447.
17. Zinder, J.C. and C.D. Lima, *Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors*. Genes Dev, 2017. **31**(2): p. 88-100.
18. Morton, D.J., et al., *The RNA Exosome and RNA Exosome-linked Disease*. RNA, 2017.
19. Schneider, C. and D. Tollervey, *Threading the barrel of the RNA exosome*. Trends in Biochemical Sciences, 2013. **38**(10): p. 485-493.
20. Schoenberg, D.R. and L.E. Maquat, *Regulation of cytoplasmic mRNA decay*. Nat Rev Genet, 2012. **13**(4): p. 246-59.
21. Houseley, J. and D. Tollervey, *The many pathways of RNA degradation*. Cell, 2009. **136**(4): p. 763-76.
22. Zhouravleva, G. and S. Bondarev, *Gene Duplication and the Origin of Translation Factors*. 2011.
23. Maquat, L.E., *Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics*. Nature Reviews Molecular Cell Biology, 2004. **5**(2): p. 89-99.
24. Fasken, M.B., et al., *The RNA Exosome and Human Disease*. Methods Mol Biol, 2020. **2062**: p. 3-33.
25. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nat Genet, 2012. **44**(6): p. 704-8.
26. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. Nat Commun, 2014. **5**: p. 4287.
27. Burns, D.T., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. Neuromuscul Disord, 2017. **27**: p. S38.
28. Di Donato, N., et al., *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, 2016. **53**(6): p. 419-25.

29. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. Hum Mol Genet, 2020. **29**(13): p. 2218-2239.
30. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
31. Somashekar, P.H., et al., *Bi-allelic missense variant, p.Ser35Leu in EXOSC1 is associated with pontocerebellar hypoplasia*. Clin Genet, 2021. **99**(4): p. 594-600.
32. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 227-239.
33. Morton, D.J., et al., *The RNA exosome and RNA exosome-linked disease*. Rna, 2018. **24**(2): p. 127-142.
34. Schneider, C., et al., *Transcriptome-wide Analysis of Exosome Targets*. Molecular Cell, 2012. **48**(3): p. 422-433.
35. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases*. Cell, 1997. **91**(4): p. 457-466.
36. Mitchell, P., et al., *The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases*. Cell, 1997. **91**(4): p. 457-66.
37. Chekanova, J.A., et al., *Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome*. Cell, 2007. **131**(7): p. 1340-53.
38. Gudipati, R.K., et al., *Extensive degradation of RNA precursors by the exosome in wild-type cells*. Mol Cell, 2012. **48**(3): p. 409-21.
39. Pefanis, E., et al., *Noncoding RNA transcription targets AID to divergently transcribed loci in B cells*. Nature, 2014. **514**(7522): p. 389-93.
40. Allmang, C., et al., *Functions of the exosome in rRNA, snoRNA and snRNA synthesis*. Embo Journal, 1999. **18**(19): p. 5399-5410.
41. van Hoof, A., P. Lennertz, and R. Parker, *Yeast Exosome Mutants Accumulate 3'-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs*. Molecular and Cellular Biology, 2000. **20**(2): p. 441-452.
42. Wyers, F., et al., *Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase*. Cell, 2005. **121**(5): p. 725-737.
43. Bonneau, F., et al., *The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation*. Cell, 2009. **139**(3): p. 547-59.

44. Liu, Q., J.C. Greimann, and C.D. Lima, *Reconstitution, activities, and structure of the eukaryotic RNA exosome*. Cell, 2006. **127**(6): p. 1223-1237.
45. Makino, D.L., M. Baumgaertner, and E. Conti, *Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex*. Nature, 2013. **495**(7439): p. 70-75.
46. Wasmuth, E.V., K. Januszyk, and C.D. Lima, *Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA*. Nature, 2014. **511**(7510): p. 435-9.
47. Zinder, J.C., E.V. Wasmuth, and C.D. Lima, *Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3*. Mol Cell, 2016. **64**(4): p. 734-745.
48. Lorentzen, E., et al., *RNA channelling by the archaeal exosome*. Embo Reports, 2007. **8**(5): p. 470-476.
49. Büttner, K., K. Wenig, and K.-P. Hopfner, *Structural Framework for the Mechanism of Archaeal Exosomes in RNA Processing*. Molecular Cell, 2005. **20**(3): p. 461-471.
50. Navarro, M.V.A.S., et al., *Insights into the Mechanism of Progressive RNA Degradation by the Archaeal Exosome* ^{*} *</sup>. Journal of Biological Chemistry, 2008. **283**(20): p. 14120-14131.*
51. Lorentzen, E. and E. Conti, *Structural Basis of 3' End RNA Recognition and Exoribonucleolytic Cleavage by an Exosome RNase PH Core*. Molecular Cell, 2005. **20**(3): p. 473-481.
52. Lorentzen, E. and E. Conti, *Crystal Structure of a 9-Subunit Archaeal Exosome in Pre-Catalytic States of the Phosphorolytic Reaction*. Archaea, 2012. **2012**: p. 721869.
53. Lorentzen, E., et al., *The archaeal exosome core is a hexameric ring structure with three catalytic subunits*. Nature Structural & Molecular Biology, 2005. **12**(7): p. 575-581.
54. Lu, C., F. Ding, and A. Ke, *Crystal Structure of the S. solfataricus Archaeal Exosome Reveals Conformational Flexibility in the RNA-Binding Ring*. PLOS ONE, 2010. **5**(1): p. e8739.
55. Hou, D., M. Ruiz, and E.D. Andrulis, *The ribonuclease Dis3 is an essential regulator of the developmental transcriptome*. BMC Genomics, 2012. **13**.
56. Lim, S.J., et al., *Genome-wide localization of exosome components to active promoters and chromatin insulators in Drosophila*. Nucleic Acids Res, 2013. **41**(5): p. 2963-80.
57. van Hoof, A., P. Lennertz, and R. Parker, *Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs*. Mol Cell Biol, 2000. **20**(2): p. 441-52.

58. van Hoof, A., et al., *Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA*. Mol Cell Biol, 2000. **20**(21): p. 8230-43.
59. Fernández-Pevida, A., D. Kressler, and J. de la Cruz, *Processing of preribosomal RNA in Saccharomyces cerevisiae*. Wiley Interdiscip Rev RNA, 2015. **6**(2): p. 191-209.
60. Henras, A.K., et al., *An overview of pre-ribosomal RNA processing in eukaryotes*. Wiley Interdiscip Rev RNA, 2015. **6**(2): p. 225-42.
61. Hongay, C.F., et al., *Antisense transcription controls cell fate in Saccharomyces cerevisiae*. Cell, 2006. **127**(4): p. 735-745.
62. Lloret-Llinares, M., et al., *Relationships between PROMPT and gene expression*. RNA biology, 2016. **13**(1): p. 6-14.
63. Martens, J.A., P.-Y.J. Wu, and F. Winston, *Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae*. Genes & Development, 2005. **19**(22): p. 2695-2704.
64. Uhler, J.P., C. Hertel, and J.Q. Svejstrup, *A role for noncoding transcription in activation of the yeast <i>PHO5</i> gene*. Proceedings of the National Academy of Sciences, 2007. **104**(19): p. 8011-8016.
65. Zinder, J.C. and C.D. Lima, *Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors*. Genes & Development, 2017. **31**(2): p. 88-100.
66. Ogami, K., Y. Chen, and J.L. Manley, *RNA surveillance by the nuclear RNA exosome: mechanisms and significance*. Non-coding RNA, 2018. **4**(1): p. 8.
67. Briggs, M.W., K.T. Burkard, and J.S. Butler, *Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation*. J Biol Chem, 1998. **273**(21): p. 13255-63.
68. Mitchell, P., et al., *Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs*. Mol Cell Biol, 2003. **23**(19): p. 6982-92.
69. Milligan, L., et al., *A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts*. Mol Cell Biol, 2008. **28**(17): p. 5446-57.
70. Schilders, G., et al., *MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation*. Nucleic Acids Research, 2005. **33**(21): p. 6795-6804.
71. van Hoof, A., P. Lennertz, and R. Parker, *Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs*. Molecular and cellular biology, 2000. **20**(2): p. 441-452.

72. Liang, S., et al., *A DEAD-box-family protein is required for nucleocytoplasmic transport of yeast mRNA*. Mol Cell Biol, 1996. **16**(9): p. 5139-46.
73. de la Cruz, J., et al., *Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae*. Embo j, 1998. **17**(4): p. 1128-40.
74. Taylor, L.L., et al., *The Mtr4 ratchet helix and arch domain both function to promote RNA unwinding*. Nucleic acids research, 2014. **42**(22): p. 13861-13872.
75. Schuch, B., et al., *The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase*. Embo Journal, 2014. **33**(23): p. 2829-2846.
76. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. Cell Reports, 2017. **20**(10): p. 2279-2286.
77. Wasmuth, E.V., et al., *Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase*. Elife, 2017. **6**: p. 24.
78. Schuller, J.M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. Science, 2018. **360**(6385): p. 219-222.
79. Weick, E.M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. Cell, 2018. **173**(7): p. 1663-1677 e21.
80. Weir, J.R., et al., *Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance*. Proc Natl Acad Sci U S A, 2010. **107**(27): p. 12139-44.
81. Stuparevic, I., et al., *Cotranscriptional Recruitment of RNA Exosome Cofactors Rrp47p and Mpp6p and Two Distinct Trf-Air-Mtr4 Polyadenylation (TRAMP) Complexes Assists the Exonuclease Rrp6p in the Targeting and Degradation of an Aberrant Messenger Ribonucleoprotein Particle (mRNP) in Yeast*. Journal of Biological Chemistry, 2013. **288**(44): p. 31816-31829.
82. Falk, S., et al., *The molecular architecture of the TRAMP complex reveals the organization and interplay of its two catalytic activities*. Mol Cell, 2014. **55**(6): p. 856-867.
83. Houseley, J. and D. Tollervey, *The nuclear RNA surveillance machinery: the link between ncRNAs and genome structure in budding yeast?* Biochim Biophys Acta, 2008. **1779**(4): p. 239-46.
84. Lubas, M., et al., *Interaction profiling identifies the human nuclear exosome targeting complex*. Mol Cell, 2011. **43**(4): p. 624-37.

85. Staals, R.H., et al., *Dis3-like 1: a novel exoribonuclease associated with the human exosome*. *Embo j*, 2010. **29**(14): p. 2358-67.
86. Lohr, J.G., et al., *Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy*. *Cancer Cell*, 2014. **25**(1): p. 91-101.
87. Chapman, M.A., et al., *Initial genome sequencing and analysis of multiple myeloma*. *Nature*, 2011. **471**(7339): p. 467-72.
88. Alexander, D.D., et al., *Multiple myeloma: a review of the epidemiologic literature*. *Int J Cancer*, 2007. **120 Suppl 12**: p. 40-61.
89. Laubach, J., P. Richardson, and K. Anderson, *Multiple myeloma*. *Annu Rev Med*, 2011. **62**: p. 249-64.
90. Weissbach, S., et al., *The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma*. *Br J Haematol*, 2015. **169**(1): p. 57-70.
91. Tomecki, R., et al., *Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target*. *Nucleic Acids Res*, 2014. **42**(2): p. 1270-90.
92. Boyle, E.M., et al., *BRAF and DIS3 Mutations Associate with Adverse Outcome in a Long-term Follow-up of Patients with Multiple Myeloma*. *Clin Cancer Res*, 2020. **26**(10): p. 2422-2432.
93. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. *J Neurol*, 2013. **260**(7): p. 1866-70.
94. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. *Nature Communications*, 2014. **5**.
95. Di Donato, N., et al., *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*. *Journal of Medical Genetics*, 2016. **53**(6): p. 419-425.
96. Eggens, V.R.C., et al., *EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations*. *Orphanet Journal of Rare Diseases*, 2014. **9**(1): p. 23.
97. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
98. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. *Nature Genetics*, 2012. **44**(6): p. 704-U134.

99. Burns, D., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. *Neuromuscul Disord*, 2017. **27**: p. S38.
100. Yang, X., et al., *Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis*. *Human Molecular Genetics*, 2019. **29**(4): p. 541-553.
101. Schnell, H.M., et al., *Chaperone-mediated assembly of the proteasome core particle – recent developments and structural insights*. *Journal of Cell Science*, 2022. **135**(8): p. jcs259622.
102. Schnell, H.M., et al., *Structures of chaperone-associated assembly intermediates reveal coordinated mechanisms of proteasome biogenesis*. *Nature Structural & Molecular Biology*, 2021. **28**(5): p. 418-425.
103. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. *Genetics*, 2017. **205**(1): p. 221-+.
104. Gillespie, A., et al., *Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in *S. cerevisiae**. *RNA*, 2017. **23**(4): p. 466-472.
105. Han, J. and A. van Hoof, *The RNA Exosome Channeling and Direct Access Conformations Have Distinct In Vivo Functions*. *Cell Reports*, 2016. **16**(12): p. 3348-3358.
106. Tollervey, D. and T. Kiss, *Function and synthesis of small nucleolar RNAs*. *Curr Opin Cell Biol*, 1997. **9**(3): p. 337-42.
107. van Hoof, A., P. Lennertz, and R. Parker, *Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast*. *EMBO J*, 2000. **19**(6): p. 1357-65.
108. Sterrett, M. and A. Corbett, *Post-Transcriptional Regulation of Gene Expression and Human Genetic Disease*. 2019. p. 1-7.
109. Schuller Jan, M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. *Science*, 2018. **360**(6385): p. 219-222.

Chapter II: A Budding Yeast Model for Human Disease Mutations in the EXOSC2 Cap Subunit of the RNA Exosome Complex

The research and data presented in this chapter is published in *RNA* as “A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex”

Sterrett, M. C.*, L. Enyenihi,* S. W. Leung, L. Hess, S. E. Strassler, D. Farchi, R. S. Lee, E. S. Withers, I. Kremsky, R. E. Baker, M. A. Basrai, A. van Hoof, M. B. Fasken and A. H. Corbett (2021). "A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex." *Rna* **27**(9): 1046-1067.

* Indicate co-authorship

Author Contributions:

M.C Sterrett created all the data for published figures with assistance from L. Enyenihi. **M.C. Sterrett** also compiled and wrote the manuscript. D. Farchi, R.S. Lee, E.S. Withers modeled amino acid substitutions in solved structures under the guidance of S. E. Strassler. I. Kremsky and M.B. Fasken performed RNA-Seq analysis. L. Enyenihi and D. Farchi were undergraduate mentees of **M.C.Sterrett**.

2.1 Abstract

RNA exosomopathies, a growing family of diseases, are linked to missense mutations in genes encoding structural subunits of the evolutionarily conserved, 10-subunit exoribonuclease complex, the RNA exosome. This complex consists of a 3-subunit cap, a 6-subunit, barrel-shaped, core and a catalytic base subunit. While a number of mutations in RNA exosome genes cause pontocerebellar hypoplasia, mutations in the cap subunit gene *EXOSC2* cause an apparently distinct clinical presentation that has been defined as a novel syndrome SHRF (Short stature, Hearing loss, Retinitis pigmentosa and distinctive Facies). We generated the first *in vivo* model of the SHRF pathogenic amino acid substitutions using budding yeast by modeling pathogenic *EXOSC2* missense mutations (p.Gly30Val and p.Gly198Asp) in the orthologous *S. cerevisiae* gene *RRP4*. The resulting *rrp4* mutant cells show defects in cell growth and RNA exosome function. Consistent with altered RNA exosome function, we detect significant transcriptomic changes in both coding and non-coding RNAs in *rrp4-G226D* cells that model *EXOSC2* p.Gly198Asp, suggesting defects in nuclear surveillance. Biochemical and genetic analyses suggest that the Rrp4 G226D variant subunit shows impaired interactions with key RNA exosome cofactors that modulate the function of the complex. These results provide the first *in vivo* evidence that pathogenic missense mutations present in *EXOSC2* impair the function of the RNA exosome. This study also sets the stage to compare exosomopathy models to understand how defects in RNA exosome function underlie distinct pathologies.

2.2 Introduction

The RNA exosome is an evolutionarily conserved, multi-subunit riboexonuclease complex that plays multiple roles in RNA processing and decay. First identified in *Saccharomyces cerevisiae* in a screen for ribosomal RNA processing (*rrp*) mutants [1, 2], the RNA exosome is essential in all systems studied thus far [1, 3-6]. In addition to critical roles in generating mature ribosomal RNA, the RNA exosome processes a variety of small non-coding RNAs (ncRNAs), including small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and transfer RNAs (tRNAs) [7-10]. Beyond processing numerous RNAs, the RNA exosome is also required for RNA decay and surveillance, including nuclear degradation of cryptic unstable transcripts (CUTs) in budding yeast that result from pervasive transcription [11-13] and cytoplasmic RNA turnover of aberrant and non-functional mRNAs [14]. Therefore, the RNA exosome has many far-reaching roles *in vivo* that affect nearly every class of RNA [12, 15].

This essential RNA processing/degradation machine is composed of nine structural subunits associated with a catalytic 3'-5' exo/endoribonuclease subunit [DIS3/DIS3L (human); Dis3/Rrp44 (budding yeast)] [1, 16]. As illustrated in Figure 1A, the 9-subunit structural barrel is composed of an upper ring of three S1/KH cap subunits (EXOSC1/2/3; Csl4/Rrp4/Rrp40) and a lower ring of six PH-like subunits (EXOSC4/7/8/9/5/6; Rrp41/Rrp42/Rrp43/Rrp45/Rrp46/Mtr3). Within the nucleus, an additional 3'-5' exonuclease subunit termed EXOSC10/Rrp6 is also associated with the complex [17, 18]. Structural studies revealed conservation in the structural organization of the RNA exosome (Figure 1B) [16, 18-21], suggesting evolutionary conservation not just of subunit sequence but of overall complex structure and organization.

As described above, a feature of the RNA exosome is the ability to both process and degrade numerous RNA targets. The specificity for this broad set of target transcripts is achieved, at least in part, through RNA exosome cofactors that associate with the complex via interactions with multiple subunits [8, 22]. These cofactors are localized to both the nucleus and the cytoplasm, raising the possibility that

these interactions facilitate targeting of distinct target transcripts in different cellular compartments. Nuclear RNA exosome cofactors have been extensively characterized in the budding yeast system, including the Rrp6 obligate binding partner Rrp47, the essential RNA helicase Mtr4, and Mpp6 [15, 22-24]. Structural studies of both the budding yeast and mammalian RNA exosome suggest Mpp6/MPH6, Rrp6/EXOSC10 and Rrp47/C1D can form composite sites that facilitate interactions between the complex and other cofactors, including Mtr4/MTR4/MTREX [25-29]. Mtr4/MTR4/MTREX aids the RNA exosome in targeting and processing target RNA, such as the 5.8S rRNA precursor (7S rRNA), and can act as part of the budding yeast TRAMP (Trf4/5-Air1/2-Mtr4 Polyadenylation) complex or the mammalian NEXT (Nuclear Exosome Targeting) complex, both of which facilitate nuclear RNA surveillance and quality control of ncRNA by the RNA exosome [22-26, 30-36]. Thus, interactions with these nuclear cofactors influence the function of the RNA exosome *in vivo*.

Recent studies have linked missense mutations in *EXOSC* genes encoding the structural subunits of the RNA exosome to various human pathologies, which comprise a growing family of diseases termed “RNA exosomopathies” [8, 37-44]. Intriguingly, these single amino acid substitutions often occur in highly conserved domains of the RNA exosome subunits. Mutations in the cap subunit gene *EXOSC3* and the core subunit gene *EXOSC8* cause forms of pontocerebellar hypoplasia (PCH1b and PCH1c, respectively), neurological disorders characterized by atrophy of the pons and cerebellum [37, 38, 41-43, 45], while mutations in the core subunit genes *EXOSC5* and *EXOSC9* have been linked to similar neurological defects including cerebellar degeneration, neuronopathy and neurodevelopmental delays [39, 44, 46]. In contrast to the other exosomopathy mutations described thus far, missense mutations in the cap subunit gene *EXOSC2* have been linked to a novel, complex syndrome characterized by retinitis pigmentosa, progressive hearing loss, premature aging, short stature, mild intellectual disability and distinctive gestalt [40], later named SHRF (Short stature, Hearing loss, Retinitis pigmentosa and distinctive Facies) (OMIM #617763) [47]. While SHRF patients do show some cerebellar atrophy [40],

the disease phenotype is distinct from PCH as well as the other neurological deficits observed in patients with other exosomopathies, suggesting a unique molecular pathology linked to *EXOSC2* mutations.

Whole exome sequencing of the three identified SHRF patients, representing two related patients and one unrelated patient, identified missense mutations in the *EXOSC2* gene that alter conserved amino acids in this cap subunit, shown in Figure 1C [40]. The two related patients have a homozygous missense mutation *EXOSC2* p.Gly30Val (G30V) in the N-terminal domain of *EXOSC2* [40]. The other patient carries compound heterozygous missense mutations *EXOSC2* p.Gly30Val and *EXOSC2* p.Gly198Asp (G30V/G198D), with the G198D missense mutation located within the K-homology RNA binding domain [40, 47]. These amino acid substitutions occur in highly conserved residues of *EXOSC2*, which are conserved across *EXOSC2*/Rrp4 orthologs from different eukaryotic species and conserved between *EXOSC2* and the *EXOSC3*/Rrp40 cap subunits of the eukaryotic RNA exosome (Figure S1). Notably, *EXOSC2* Gly30 and *EXOSC3* Gly31, an amino acid that is substituted in PCH1b patients [43], are conserved and in the same position in the two cap subunits, falling within a conserved “VxPG” consensus sequence (Figure S1). *EXOSC2* Gly198 and *EXOSC3* Trp238, another amino acid that is substituted in PCH1b patients [43], lie in the KH domains of the two cap subunits, falling within or adjacent to a conserved “GxNG” motif. The “GxNG” motif is unique to the KH domain of these RNA exosome cap subunits and is predicted to play a structural role [48]. However, when *EXOSC2* Gly30, *EXOSC2* Gly198 and *EXOSC3* Gly31, *EXOSC3* Trp238 are substituted, they give rise to distinct disease phenotypes, suggesting that similar missense mutations in *EXOSC2* and *EXOSC3* have different mechanistic effects on the function of the RNA exosome *in vivo*. Therefore, to better understand the molecular pathology of these exosomopathies, including SHRF, it is necessary to investigate the molecular and functional consequences of pathogenic amino acid substitutions that underlie each disease.

A previous study provided some important insights into how the *EXOSC2* mutations that cause SHRF could contribute to pathology [47]. This study employed several different approaches, including using patient B-lymphoblasts, *in vitro* cell culture and a *D. melanogaster* model depleted for the fly

EXOSC2/Rrp4 ortholog. Taken together, results from this study suggest that EXOSC2 dysfunction could compromise downstream molecular pathways, including neurodevelopment and autophagy [47]. While informative in probing the molecular pathology that may underlie the SHRF syndrome, a limitation of this study is that the authors did not examine known targets of the RNA exosome nor did they examine the *in vivo* consequences of the SHRF-linked EXOSC2 variants within a whole organism. Given that this diverse class of RNA exosomopathies arises from amino acid substitutions in structural subunits of a singular complex, assessing defects in RNA exosome function *in vivo* is critical for a holistic understanding of the molecular and functional consequences underlying each disease phenotype. Previous studies have assessed the functional and molecular consequences of exosomopathy-linked *EXOSC3* and *EXOSC5* mutations *in vivo* using yeast and fly genetic model systems [44, 49-52]. Utilizing a genetic model system to explore the consequences of the specific amino acid changes that occur in SHRF can provide insight into how RNA exosome function is altered in disease.

To explore the functional consequences of the amino acid substitutions in EXOSC2 that occur in SHRF, we took advantage of the budding yeast model system. We generated variants of the *S. cerevisiae* EXOSC2 ortholog, Rrp4, that model the pathogenic amino acid substitutions and examined their function in budding yeast. Our results show that the yeast Rrp4 G58V variant, corresponding to the EXOSC2 G30V variant, is not able to replace the function of the essential *RRP4* gene. In contrast, cells that express the Rrp4 G226D variant, corresponding to the EXOSC2 G198D variant, show impaired cell growth and defects in RNA exosome function, revealing that this Rrp4 G226D variant is functional but impaired. Based on RNA-Seq analysis, the *rrp4-G226D* cells show broad transcriptomic changes with defects particularly in nuclear surveillance by the RNA exosome. Genetic and biochemical studies demonstrate that the *rrp4-G226D* mutant exhibits negative genetic interactions with RNA exosome cofactor mutants, likely suggesting defects in association with Mtr4. Combined, these results suggest that amino acid changes in Rrp4 that model those in EXOSC2 impair the overall function of the RNA exosome, potentially through impairment of interactions with the exosome cofactor Mtr4.

2.3 Materials and Methods

Chemicals and media

All chemicals were obtained from Sigma-Aldrich (St. Louis, MO), United States Biological (Swampscott, MA), or Fisher Scientific (Pittsburgh, PA) unless otherwise noted. All media were prepared by standard procedures [74].

Protein structure analysis

We used the cryo-EM structure (PDB 6D6R) of the human nuclear RNA exosome at 3.45Å resolution [28] and the cryo-EM structure (PDB 6FSZ) of the budding yeast nuclear RNA exosome at 4.6Å [29]. Structural modeling was performed using the PyMOL viewer (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) [75]. The mCSM-PP12 [76], Polymorphism Phenotyping V2 (PolyPhen-2) [77], Protein Variation Effect Analyzer (PROVEAN) [78, 79] and SNAP-2 [80] webservers were used for predicting the effects of the *EXOSC2* mutations on protein stability and function.

***Saccharomyces cerevisiae* strains and plasmids**

All DNA manipulations were performed according to standard procedures [81]. *S. cerevisiae* strains and plasmids used in this study are listed in Table S1. The *rrp4Δ* (yAV1103) and *rrp40Δ* (yAV1107) strains were previously described [82, 83]. The *RRP43-TAP* (ACY2788) strain was obtained from Horizon Discovery Biosciences Limited and was previously described [84]. The *RRP43-TAP rrp4Δ* (ACY2803) strain was constructed by deletion of the genomic *RRP4* ORF in the *RRP43-TAP* (ACY2788) strain containing a *RRP4 URA3* maintenance plasmid by homologous recombination using a *RRP4-UTR neoMX4* PCR product. The *rrp4Δ mpp6Δ* (ACY2471), *rrp4Δ rrp47Δ* (ACY2474), *rrp4Δ rrp6Δ* (ACY2478) strains and the *rrp40Δ mpp6Δ* (ACY2638), *rrp40Δ rrp47Δ* (ACY2462), *rrp40Δ rrp6Δ* (ACY2466) strains were constructed by deletion of the *MPP6*, *RRP47*, and *RRP6* ORF in the *rrp4Δ* (yAV1103) and *rrp40Δ* (yAV1107) strains by homologous recombination using *MPP6-*, *RRP47-*, or

RRP6-UTR natMX4 PCR products. The *rrp4Δ mtr4Δ* (ACY2536) and *rrp40Δ mtr4Δ* (ACY2540) strains were constructed by consecutive deletion of the genomic *MTR4* ORF and *RRP4* ORF or *RRP40* ORF in a wild-type (W303) strain harboring a [-*MTR4*; *RRP4*; *URA3*] (pAC3714) or [*MTR4*; *RRP40*; *URA3*] (pAC3713) maintenance plasmid by homologous recombination using *MTR4-UTR natMX4* and *RRP4-UTR* or *RRP40-UTRneoMX4* PCR products. Construction of *RRP40-2xMyc* and *rrp40-2xMyc* variant plasmids (pAC3161, pAC3162 and pAC3259) was reported previously [49]. The *RRP4-2xMyc LEU2 CEN6* (pAC3474) plasmid was constructed by PCR amplification of the endogenous promoter, 5'-UTR and ORF of the *RRP4* gene from *S. cerevisiae* genomic DNA and cloning into pRS315 plasmid containing a C-terminal 2xMyc tag and the *ADHI* 3'-UTR [53]. The *rrp4-G58V-2xMyc* (pAC3476) and *rrp4-G226D-2xMyc* (pAC3477) plasmids were generated by site-directed mutagenesis of the *RRP4-2xMyc* (pAC3474) plasmid using oligonucleotides containing the SHRF syndrome-linked G58V and G226D missense mutations and the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The untagged *RRP4/rrp4-G226D* (pAC3656, pAC3659) and *RRP40/rrp40-W195R* (pAC3652, pAC3655) plasmids and Myc-tagged *RRP4/rrp4-G226D* (pAC3669, pAC3672) plasmid containing native 3'-UTRs were generated by excision of the *2xMyc-ADHI* 3'-UTR from each *RRP4/40-Myc LEU2 CEN6* plasmid by restriction digestion and cloning of the native *RRP4* 3'-UTR, *RRP40* 3'-UTR, or *2xMyc-RRP4* 3'-UTR PCR product into the relevant plasmid using NEBuilder HiFi Assembly (New England BioLabs). The *MTR4 HIS3 CEN6* (pAC4096) plasmid was constructed by PCR amplification of the endogenous promoter, 5'-UTR, ORF, and 3'UTR of the *MTR4* gene from *S. cerevisiae* genomic DNA and cloning into pRS313 [53]. The *mtr4-F7A-F10A* (pAC4099) plasmid was generated by site-directed mutagenesis of the *MTR4 HIS CEN6* plasmid (pAC4096) using oligonucleotides containing the F7A and F10A missense mutations and the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The *MTR4-2xFLAG* (pAC3719) plasmid was constructed by PCR amplification of the *MTR4* promoter-5'-UTR-ORF(No Stop) and *2xFLAG-native MTR4* 3'-UTR using *MTR4* template plasmid [pAC2897;[85]] and oligonucleotides encoding 2xFLAG epitopes and cloning into the pRS313 plasmid [53].

***S. cerevisiae* transformations and growth assays**

All yeast transformations were performed according to the standard Lithium Acetate (LiOAc) protocol. Cells were grown overnight to saturation in a 30°C incubator in liquid YEPD (1% yeast extract, 2% peptone, 2% dextrose, in distilled water). Cell concentrations were normalized to an OD₆₀₀ = 0.4 in 10 mL YEPD and incubated at 30°C for 5 hr. The cells were washed with TE/LiOAc and resuspended in TE/LiOAc to a concentration of 2 x 10⁹ cells/mL. To the cells (50 µL), plasmid DNA, single-stranded carrier DNA (5 µL), and PEG/TE/LiOAc (300 µL) were added and cells were agitated for 30 mins at 30°C. DMSO (35 µL) was added and the cells were heat shocked at 42°C for 15 min. The cells were washed and plated onto selective media.

To test the *in vivo* function of the *rrp4* variants that model the *EXOSC2* variants in SHRF syndrome, a standard plasmid shuffle assay was employed. The *rrp4*Δ (yAV1103) cells containing a *RRP4 URA3* maintenance plasmid and transformed with vector (pRS315), *RRP4-2xMyc* (pAC3474), *rrp4-G8A-2xMyc* (pAC3476), or *rrp4-G226D-2xMyc* (pAC3477) plasmid were grown overnight and serially diluted and spotted onto Ura⁻ Leu⁻ minimal media plates, which select for cells that contain both the *RRP4 URA3* maintenance plasmid and the *RRP4/rrp4 LEU2* plasmid, and 5-FOA Leu⁻ minimal media plates, which select for cells that lack the *RRP4 URA3* maintenance plasmid and contain only the *RRP4/rrp4 LEU2* plasmid. The plates were incubated at 30°C and 37°C for 2 days.

The *in vivo* function of the *rrp4-G226D* variant and the *rrp40-W195R* variant was assessed in growth assays on solid media and in liquid culture. For growth on solid media, *rrp4*Δ (yAV1103) cells containing only *RRP4* (pAC3656) or *rrp4-G226D* (pAC3659) and *rrp40*Δ (yAV1107) cells containing only *RRP40* (pAC3652) or *rrp40-W195R* (pAC3655) were grown in 2mL Leu⁻ minimal media overnight at 30°C to saturation. Cell concentrations were normalized to an OD₆₀₀ = 0.5, serially diluted in 10-fold dilutions, spotted onto Leu⁻ minimal media plates, and grown at 25°C, 30°C, and 37°C for 2-3 days. For

growth in liquid culture, cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation, diluted to an OD₆₀₀ = 0.01 in Leu⁻ minimal media in a 24-well plate, and growth at 37°C was monitored and recorded at OD₆₀₀ in a BioTek® SynergyMx microplate reader with Gen5™ v2.04 software over 24 hr. Technical triplicates of each strain were measured, and the average of these triplicates was calculated and graphed.

Immunoblotting

For analysis of C-terminally Myc-tagged Rrp4 protein expression levels, *rrp4Δ* (yAV1103) cells expressing only Rrp4-2xMyc (pAC3669) or *rrp4-G226D-2xMyc* (pAC3672) were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation and 10 mL cultures with an OD₆₀₀ = 0.2 were prepared and grown at 30°C and 37°C for 5 hr. Additionally, *rrp4Δ* (yAV1103) cells containing *RRP4 URA3* maintenance plasmid and expressing Rrp4-2xMyc (pAC3474), *rrp4-G58V* (pAC3476), or *rrp4-G226D-2xMyc* (pAC3477) were grown in 2 mL Ura⁻ Leu⁻ minimal media overnight at 30°C and 10 mL cultures with an OD₆₀₀ = 0.2 were prepared and grown at 30°C for 5 hr. Cell pellets were collected by centrifugation, transferred to 2 mL screw-cap tubes and stored at -80°C. Yeast cell lysates were prepared by resuspending cell pellets in 0.3-0.5 mL of RIPA-2 Buffer (50 mM Tris-HCl, pH 8; 150 mM NaCl; 0.5% sodium deoxycholate; 1% NP40; 0.1% SDS) supplemented with protease inhibitors [1 mM PMSF; Pierce™ Protease Inhibitors (Thermo Fisher Scientific)], and 300 μL of glass beads. Cells were disrupted in a Mini Bead Beater 16 Cell Disrupter (Biospec) for 4 × 1 min at 25°C with 1 min on ice between repetitions, and then centrifuged at 16,000 × g for 15 min at 4°C. Protein lysate concentration was determined by Pierce BCA Protein Assay Kit (Life Technologies). Whole cell lysate protein samples (40 μg) in reducing sample buffer (50 mM Tris HCl, pH 6.8; 100 mM DTT; 2% SDS; 0.1% Bromophenol Blue; 10% Glycerol) were resolved on 4–20% Criterion™ TGX Stain-Free™ precast polyacrylamide gels (Bio-Rad), transferred to nitrocellulose membranes (Bio-Rad) and Myc-tagged Rrp4 proteins were detected with anti-Myc monoclonal antibody 9B11 (1:2000; Cell Signaling). The 3-phosphoglycerate kinase (Pgc1) protein was detected using anti-Pgc1 monoclonal antibody (1:30,000; Invitrogen) as a

loading control.

Quantitation of immunoblotting

The protein band intensities/areas from immunoblots were quantitated using ImageJ v1.4 software (National Institute of Health, MD; <http://rsb.info.nih.gov/ij/>) and mean fold changes in protein levels were calculated in Microsoft Excel (Microsoft Corporation). To quantitate the mean fold change in Rrp4 G226D-Myc variant level relative to wild-type Rrp4-Myc level in *rrp4* Δ cells grown at 30°C and 37°C from three immunoblots (Figure 3D) or the fold change in *rrp4*-G58V-Myc and *rrp4*-G226D-Myc level in *rrp4* Δ cells containing untagged *RRP4* from three immunoblots (Figure 3E), R/*rrp4*-Myc intensity was first normalized to loading control Pgk1 intensity and then normalized to wildtype Rrp4-Myc intensity at 30°C or 37°C for each immunoblot. The mean fold change in R/*rrp4* -Myc level relative to Rrp4-Myc and standard error of the mean are graphed in Figure 3F and 3G. To quantitate the mean percent bound of Mtr4-FLAG co-immunoprecipitated with Rrp4-Myc or Rrp4 G226D-Myc (Figure 9C), background intensity of Mtr4-FLAG was subtracted from the bound Mtr4-Flag intensity. Then bound and input Mtr4-FLAG intensity was normalized to input Pgk1 intensity. Percent bound was calculated by dividing normalized bound Mtr4-FLAG by normalized input Mtr4-FLAG. To quantitate the mean percent bound Rrp4-Myc or Rrp4 G226D-Myc that immunoprecipitates (Figure 9C), bound and input R/*rrp4*-Myc was normalized to input Pgk1 intensity. Percent bound was calculated by dividing normalized bound R/*rrp4*-Myc intensity by normalized input R/*rrp4*-Myc intensity. Mean percent bound Mtr4-FLAG and R/*rrp4*-Myc and standard error of the mean are graphed in Figure 9D and 9E.

Northern blotting

For analysis of 5.8S pre-rRNA processing - detection of 7S pre-rRNA and processing intermediates - in *rrp4* and *rrp40* mutant cells, *rrp4* Δ (yAV1103) cells containing *RRP4-2xMyc* (pAC3474) or *rrp4-G226D-2xMyc* (pAC3477) and *rrp40* Δ (yAV1107) cells containing *RRP40-2xMyc* (pAC3161), *rrp40-G8A-2xMyc* (pAC3162), or *rrp40-W195R-2xMyc* (pAC3259) were grown in 2 mL Leu- minimal media

overnight at 30°C, 10 mL cultures with an OD₆₀₀ = 0.4 were prepared and grown at 37°C for 5 hr. Cells were collected by centrifugation (2,163 x g), transferred to 2 mL screw cap tubes and stored at -80°C. Total RNA from cells was resolved on an Criterion TBE-Urea polyacrylamide gel (Bio-Rad), blotted to a nylon membrane and membrane was probed with radiolabeled 5.8S-ITS2 rRNA (boundary) oligonucleotide (AC4211/Probe 020-5'-TGAGAAGGAAATGACGCT) to detect 7S pre-rRNA and intermediates and stained with methylene blue stain to visualize 5.8S rRNA as a loading control. Total RNA (5µg) was mixed with equal volume of RNA loading dye (1xTBE; 12% Ficoll; 7M Urea; 0.01 bromophenol blue; 0.02% xylene cyanol) and resolved on 10% TBE-Urea polyacrylamide gel in 1xTBE at 200V for 1.5 hr. RNA was transferred to Hybond™-N+ nylon membrane (Amersham, GE Healthcare) at 15V for 100 min in 1xTBE and cross-linked to membrane with UV light (120,000 µJoules) using UV Stratalinker® 2400 (Stratagene). Membrane was incubated in Rapid-hyb hybridization buffer (Amersham, GE healthcare) at 37°C for 1 hr. DNA oligonucleotide (100 ng) was 5'-end labeled with [γ -P32]-ATP (PerkinElmer) using polynucleotide kinase (New England Biolabs) at 37°C for 30 min. [P32]-Labeled oligonucleotide probe was purified through G25 microspin column (GE Healthcare), heated at 100°C for 5 min, and added to hybridization buffer. Oligonucleotide probe was hybridized to membrane in hybridization buffer at 37°C overnight. Following removal of hybridization buffer, membrane was rinsed twice in 5 x SSPE; 0.1% SDS at 25°C and washed twice in 0.5 x SSPE; 0.1% SDS at 37°C for 20 min each. Membrane was exposed to phosphoscreen overnight and imaged using Typhoon FLA 7000 phosphoimager (GE Healthcare).

Total RNA Isolation

Total RNA from *S. cerevisiae rrp4* and *rrp40* mutant cells was isolated using TRIzol (Invitrogen) for qRT-PCR and northern blotting and MasterPure™ Yeast RNA Purification Kit (Epicentre, Lucigen) for RNA-Seq. *S. cerevisiae* cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation. Cultures were diluted in 10 mL Leu⁻ minimal media to an OD₆₀₀ = 0.2 and grown for 5 hr at 37°C. Cells were pelleted by centrifugation, transferred to in 2 mL screw cap tubes, and stored at -80°C. To prepare total RNA using TRIzol, cells were resuspended in 1 mL TRIzol (Invitrogen) with 300 µL of

glass beads. Cell samples were disrupted in a Biospec Mini Bead Beater 16 Cell Disrupter for 2 min at 25°C. For each sample, 100 μ L of 1-bromo-3-chloropropane (BCP) was added, sample was vortexed for 15 sec, and incubated at 25°C for 2 min. The sample was centrifuged at 16,300 \times g for 8 min at 4°C, and the upper layer was transferred to a fresh microfuge tube. RNA was precipitated with 500 μ L isopropanol and sample was vortexed for 10 sec to mix. Total RNA was pelleted by centrifugation at 16,300 \times g for 8 min at 4°C. The RNA pellet was washed with 1 mL 75% ethanol, centrifuged at 16,300 \times g for 5 min at 4°C, and air-dried for 15 min. Total RNA was resuspended in 50 μ L diethylpyrocarbonate (DEPC, Sigma)-treated water and stored at -80°C . Total RNA was prepared using MasterPure™ Yeast RNA Purification Kit (Epicentre, Lucigen) according to manufacturer's protocol. Total RNA was resuspended in 50 μ L DEPC-treated water and stored at -80°C .

qRT-PCR

All oligonucleotides used in this study are summarized in Table S2. For analysis of steady-state RNA levels using quantitative PCR, three independent biological replicates of *rrp4 Δ* (yAV1103) cells containing only *RRP4* (pAC3656) or *rrp4-G226D* (pAC3659) and *rrp40 Δ* (yAV1107) cells containing only *RRP40* (pAC3652) or *rrp40-W195R* (pAC3655) were grown in 2mL Leu⁻ minimal media overnight at 30°C, 10 mL cultures with an OD₆₀₀ = 0.2 were prepared and cells were grown at 37°C for 5 hr. Total RNA was isolated from cell pellets and 1 μ g of total RNA was reverse transcribed to first strand cDNA using the M-MLV Reverse Transcriptase (Invitrogen) according to manufacturer's protocol. Quantitative PCR was performed on technical triplicates of cDNA (10 ng) from three independent biological replicates using gene specific primers (0.5mM; Table S2), QuantiTect SYBR Green PCR master mix (Qiagen) on a StepOnePlus Real-Time PCR machine (Applied Biosystems; Tanneal=55°C, 44 cycles). *ALG9* was used as an internal control. The mean RNA levels were calculated by the $\Delta\Delta\text{Ct}$ method [86]. Mean levels of RNA calculated in mutant cells are normalized to mean levels in wild-type cells and converted and graphed as RNA fold change relative to wild-type. All primers used are summarized in Table S2.

RNA-Seq analysis

RNA-Seq was performed on three independent biological replicates of *rrp4* Δ (yAV1103) cells containing *RRP4-2xMyc* (pAC3474) or *rrp4-G226D-2xMyc* (pAC3477) as the sole copy of *RRP4* grown at 37°C. Cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C, diluted to an OD₆₀₀ = 0.4 in 10 mL Leu⁻ minimal media, grown at 37°C for 5 hr, and collected and stored at -80°C. Total RNA was isolated, rRNA was depleted, and stranded cDNA libraries were prepared using TruSeq Total RNA Stranded Library Prep kit (Illumina). Paired-end sequencing of the cDNA libraries was performed on a HiSeq4000 instrument (2 x 150 cycles) at Frederick National Laboratory for Cancer Research (FNLRCR) at the CCR Sequencing Facility, NCI, NIH, Frederick, MD. The *RRP4* samples yielded an average of 28,890,739 pass filter reads and the *rrp4-G226D* samples yielded an average of 34,644,683 pass filter reads, with a base call quality of 94% of bases with Q30 and above. The reads were mapped to the *S. cerevisiae* S288C genome assembly R64-1-1, annotated with CUTs and SUTs (Xu et al. 2009), using the STAR RNA-seq aligner [v2.7.5b (Dobin et al. 2012)]. The reads were per gene feature were counted using featureCounts [v1.6.4+galaxy2 (Liao et al. 2014)]. [69], using the STAR RNA-seq aligner [v2.7.5b [87]]. The reads per gene feature were counted using featureCounts [v1.6.4+galaxy2 [88]]. Differential gene expression analysis on raw read counts was performed using DESeq2 [v2.11.40.6+galaxy1 [89]] to identify genes significantly changed (p -value < 0.05, ≥ 1.5 fold change) in *rrp4-G226D* samples relative to *RRP4* samples. Principal component analysis (PCA) on raw read counts was also performed using DESeq2. Volcano plot of differential gene expression data was produced in Prism 8 (Graphpad Software). Piecharts of RNA class percentages in significantly altered genes were generated in Microsoft Excel for Mac (Microsoft Corp.). Gene Ontology (GO) analysis on significantly altered genes for Biological Process category was performed using the YeastEnrichr webserver [<http://amp.pharm.mssm.edu/YeastEnrichr/>] [60].

Genetic Interaction Analysis

To test genetic interactions between *rrp4-G226D* or *rrp40-W195R* and RNA exosome cofactor/subunit deletion mutants, *rrp4Δ mpp6Δ* (ACY2471), *rrp4Δ rrp47Δ* (ACY2474), and *rrp4Δ rrp6Δ* (ACY2478) cells containing only *RRP4* (pAC3656) or *rrp4-G226D* (pAC3659) and *rrp40Δ mpp6Δ* (ACY2638), *rrp40Δ rrp47Δ* (ACY2462), and *rrp40Δ rrp6Δ* (ACY2466) cells containing only *RRP40* (ACY3652) or *rrp40-W195R* (ACY3655) were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation, serially diluted, and spotted on Leu⁻ minimal media plates. The plates were incubated at 30°C and 37°C for 3 days.

To test genetic interactions between *rrp4-G226D* or *rrp40-W195R* and the *mtr4* mutant, *mtr4-F7A-F10A*, *rrp4Δ mtr4Δ* (ACY2536) cells containing the [*MTR4*; *RRP4*; *URA3*] (pAC3714) maintenance plasmid were transformed with *RRP4* (pAC3656) or *rrp4-G226D* (pAC3659) *LEU2* plasmid and *rrp40Δ mtr4Δ* (ACY2540) cells containing the [*MTR4*; *RRP40*; *URA3*] (pAC3713) maintenance plasmid were transformed with *RRP40* (pAC3652) or *rrp40-W195R* (pAC3655) *LEU2* plasmid and selected on Ura⁻ Leu⁻ minimal media plates. Transformed cells containing both the *URA3* maintenance plasmid and the *RRP4/rrp4-G226D* or *RRP40/rrp40-W195R* *LEU2* plasmid were subsequently transformed with *MTR4* (pAC4096) or *mtr4-F7A-F10A* (pAC4099) *HIS3* plasmids and selected on Ura⁻Leu⁻His⁻ minimal media plates. The transformed cells were then streaked to onto 5-FOA Leu⁻ His⁻ plates to select for cells that did not contain the *URA3* maintenance plasmid. The resulting *rrp4Δ mtr4Δ* cells containing only *RRP4* or *rrp4-G226D* *LEU2* plasmid and *MTR4* or *mtr4-F7A-F10A* *HIS3* plasmid and *rrp40Δ mtr4Δ* cells containing only *RRP40* or *rrp40-W195R* *LEU2* plasmid and *MTR4* or *mtr4-F7A-F10A* *HIS3* plasmid were grown in 2 mL Leu⁻ His⁻ minimal media overnight at 30°C to saturation, serially diluted, and spotted on Leu⁻ His⁻ minimal media plates. The plates were incubated at 30°C and 37°C for 3 days.

Co-Immunoprecipitations

All immunoprecipitations were performed using the same protocol. Cell samples were grown in 2 mL selective media overnight were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation and 10-20 mL cultures with an OD₆₀₀ = 0.2-0.3 were prepared and grown at 30°C for 5 hr. Yeast cell

lysates were prepared by resuspending cell pellets in 0.5-0.75 mL of IPP150 Buffer (10mM Tris-HCl, pH 8; 150 mM NaCl; 0.1% NP40, PMSF) supplemented with protease inhibitors [1 mM PMSF; Pierce™ Protease Inhibitors (Thermo Fisher Scientific)], and 300 μ L of glass beads. Cells were disrupted in a Mini Bead Beater 16 Cell Disrupter (Biospec) for 4-5 \times 1 min at 25°C with 1 min on ice between repetitions. Crude lysate was transferred to a chilled microcentrifuge tube and remaining beads were washed with an additional 150 μ L of IPP150 Buffer. Lysate was then cleared by centrifugation at 16,000 \times g for 10 min at 4°C. Protein lysate concentration was determined by Pierce BCA Protein Assay Kit (Life Technologies). For input samples, 40 μ g of cleared lysate was collected and frozen at -20°C. For co-immunoprecipitations, 1 mg of cleared lysate in 1 mL IPP150 Buffer was prepared, 15-30 μ L of a 1:1 bead slurry of either Pierce™ Anti-c-Myc Magnetic Beads (ThermoFisher) or IgG Sepharose® 6 Fast Flow Beads (GE Healthcare) was added, and samples were incubated at 4°C overnight with mixing. After overnight incubation, beads were washed three times in 1 mL IPP150 Buffer for 15 sec each (anti-Myc beads) or 5 min each (IgG Sepharose beads). Whole cell lysate input samples (40 μ g) and total bound samples in reducing sample buffer were boiled for 5 min® at 100°C, resolved on 4–20% Criterion™ TGX Stain-Free™ precast polyacrylamide gels (Bio-Rad), transferred to nitrocellulose membranes (Bio-Rad) and immunoblotted. Myc-tagged Rrp4 proteins were detected with mouse anti-Myc monoclonal antibody 9B11 (1:2000; Cell Signaling). TAP-tagged Rrp43 protein was detected with peroxidase anti-peroxidase (PAP) soluble complex antibody produced in rabbit (1:5000, Sigma-Aldrich). FLAG-tagged Mtr4 protein was detected with anti-DYKDDDDK (FLAG) tag rabbit monoclonal antibody D6W5B (1:2000; Cell Signaling). The 3-phosphoglycerate kinase (Pgk1) protein was detected using anti-Pgk1 monoclonal antibody (1:30,000; Invitrogen) as a loading control.

To assess association of Rrp4 G226D with the RNA exosome complex, we utilized *RRP43-TAP* (YCR035C) cells expressing *RRP4-Myc* (pAC3669), *rrp4-G58V-Myc* (pAC3670) or *rrp4-G226D-Myc* (pAC3672) and *RRP43-TAP rrp4 Δ* (ACY2803) cells expressing only *RRP4-Myc* (pAC3669) or *rrp4-G226D-Myc* (pAC3672) and immunoprecipitated Rrp43-TAP using the IgG Sepharose beads. Levels of

associated Rrp4-Myc proteins with the Rrp43-TAP tagged subunit were detected by immunoblotting. To assess association of Rrp4 G226D with the Mtr4 helicase we utilized *rrp4Δ* (yAV1103) cells expressing only *RRP4-Myc* (pAC3669) or *rrp4-G226D-Myc* (pAC3672) and co-expressing exogenous *MTR4-FLAG* (pAC3719) and immunoprecipitated Rrp4-Myc or Rrp4 G226D-Myc using anti-c-Myc beads. Levels of associated Mtr4 with the Rrp4-Myc proteins were detected by immunoblotting.

DATA DEPOSITION

The raw RNA-seq data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO) under accession GSE163106.

2.4 Results

EXOSC2 amino acid substitutions linked to SHRF are located in conserved domains

To explore how EXOSC2 G30V and EXOSC2 G198D variants could alter the structure of the EXOSC2 protein or the RNA exosome complex (Figure 1C), we modeled these EXOSC2 amino acid substitutions using recent structures of the human RNA exosome [PDB 6D6R [28]] (Figure 2A,C) and the *S. cerevisiae* RNA exosome [PDB 6FSZ [29]] (Figure 2B, D). Structural modeling shows that the EXOSC2 Gly30 residue is positioned at the interface with the core subunit EXOSC4 towards the exterior of the complex in a region with little disorder (Figure 2A). The EXOSC2 Gly30 residue lies in a β -turn next to a highly conserved proline (Pro29), which is likely essential for the region to have the flexibility needed to make the sharp turn observed in the structure. EXOSC2 Gly30 is also adjacent to an aspartic acid (Asp31), which forms a salt bridge with Arg232 of EXOSC4, likely stabilizing the interaction between the two subunits (Figure S2A). An amino acid substitution that alters the glycine at position 30 is predicted to alter the β -turn and position the Asp31 residue away from Arg232 such that the salt bridge would be disrupted. In addition, the EXOSC2 G30V substitution introduces a valine, which is

significantly larger than glycine and appears to clash with residues Asp154 and Ala191 in EXOSC4 (Figure 2A), suggesting this substitution could negatively impact the interactions between the cap EXOSC2 and core EXOSC4 subunit. In the budding yeast structure, Rrp4 Gly58, corresponding to EXOSC2 Gly30 (Figure 1C), is located at the interface with Rrp41, the budding yeast EXOSC4 ortholog, in a β -turn adjacent to a highly conserved proline (Pro57) in Rrp4 (Figure 2C), mirroring the human structural model (Figure 2A). Rrp4 Gly58 is located next to a glutamic acid (Glu59; Figure S2B) in Rrp4 that forms a salt bridge with Arg233 in Rrp41, similar to the EXOSC2-EXOSC4 interface (Figure S2A). Similar to the human structure, in the yeast exosome structure, the Rrp4 Gly58 residue is also predicted to be essential for the flexibility of the region, facilitating the β -turn, and thus stabilizing the Rrp4-Rrp41 interface (Figure S2B). Similar to the EXOSC2 G30V substitution, substitution of Gly58 in Rrp4 most likely disrupts this β -turn, disrupting the salt bridge and destabilizing the Rrp4-Rrp41 interface.

EXOSC2 Gly198 is positioned in a dense region of the subunit, surrounded by four β sheets (Figure 2C). The EXOSC2 G198D substitution introduces the large aspartic acid residue which appears to clash with neighboring residues Val85 and Asn200 (Figure 2C) and alter EXOSC2 conformation. In addition, the EXOSC2 G198D substitution introduces a polar aspartic acid residue in place of glycine with an electronegative oxygen that would undergo repulsion with the oxygen of Asn200, making the native structure depicted in Figure 2C unlikely for this variant. Structural modeling of the Rrp4 Gly226 residue, corresponding to EXOSC2 Gly198, shows that, like Gly198 in EXOSC2, Gly226 residue is positioned in a dense region of Rrp4, surrounded by four β sheets (Figure 2D). The residues neighboring Rrp4 Gly226, Val113 and Asn228, are highly conserved and correspond to EXOSC2 Val85 and Asn200, suggesting that the budding yeast Rrp4 G226D substitution can accurately model the structural changes predicted for the human EXOSC2 G198D substitution.

The online server mCSM-PPI2 was used to calculate the change in Gibbs free energy ($\Delta\Delta G$) to predict the effect of the EXOSC2 amino acid substitutions and corresponding Rrp4 amino acid substitutions on protein-protein interactions. Consistent with observations from structural modeling, the

software predicts destabilizing changes in the affinity of the protein-protein interactions for both EXOSC2 G30V ($\Delta\Delta G = -1.012$ Kcal/mol) and EXOSC2 G198D ($\Delta\Delta G = -0.509$ Kcal/mol). The EXOSC2 G198D substitution is also predicted to reduce protein stability (Score = 1.000 Polymorphism Phenotyping v2). These predictions are consistent with previous work showing that EXOSC2 G198D has reduced stability compared to wild-type EXOSC2 [47]. Furthermore, both substitutions are strongly predicted to have deleterious effects on EXOSC2 function (G30V score -7.938 and G198D score -6.35 calculated by PROVEAN; G30V score 91, G198D score 94 calculated by SNAP-2). Both Rrp4 G58V (which models EXOSC2 G30V) and Rrp4 G226D (which models EXOSC2 G198D) are predicted to decrease protein stability (Score = 1.000 Polymorphism Phenotyping v2) as well as to have deleterious effects on function (G58V score -8.981 and G226D score -6.517 calculated by PROVEAN). Rrp4 G58V is likely to alter the native protein (score 63 calculated by SNAP2), though to a slightly lower degree than calculated for the human EXOSC2 G30V variant. However, Rrp4 G226D likely results in a change to the native protein (score 92 by SNAP2), mirroring the strong effect predicted for the human EXOSC2 G198D variant. In conclusion, these *in silico* predictions (summarized in Supplementary Table S3) suggest that the pathogenic amino acid substitutions have molecular consequences that could affect RNA exosome function in both humans and budding yeast.

***Saccharomyces cerevisiae* Rrp4 variants that model the pathogenic EXOSC2 variants impair RNA exosome function**

To assess the *in vivo* consequences of the pathogenic amino acid substitutions in EXOSC2, G30V and G198D, we generated the corresponding amino acid changes in the *S. cerevisiae* ortholog Rrp4, G58V and G226D (Figure 1C). As all core RNA exosome subunits genes are essential in budding yeast [9], we first assessed whether these *rrp4* gene mutants can replace the essential *RRP4* gene. To facilitate comparison of different *rrp4* mutants, we employed a plasmid shuffle assay in which cells deleted for the genomic copy of *RRP4* are transformed with plasmids containing mutant alleles (See *Materials and Methods*). This approach ensures that the genetic background for all mutants compared to one another is

identical [53]. In this plasmid shuffle assay, *rrp4Δ* cells containing a *RRP4* maintenance plasmid and either *rrp4-G58V* or *rrp4-G226D* plasmid were serially diluted and spotted onto 5-FOA plates to select for cells that harbor the *rrp4* mutant as the sole copy of *RRP4* (Figure 3A). The *rrp4-G58V* mutant cells are not viable at any temperature tested, whereas the *rrp4-G226D* cells exhibit impaired growth defect at 37°C as compared to control *RRP4* cells (Figure 3A). Control cells expressing wild-type *RRP4* grow at all temperatures as expected. The impaired growth of *rrp4-G226D* mutant cells was further analyzed by serial dilution and spotting on solid minimal media (Figure 3B) and in a liquid media growth assay (Figure 3C). On solid media and in liquid culture, the *rrp4-G226D* cells show impaired growth at 37°C compared to control *RRP4* cells (Figure 3B, 3C). For comparison, we also assessed the growth of the previously characterized *rrp40-W195R* mutant cells [49, 51], which solely express the *rrp40-W195R* mutant corresponding to the *EXOSC3-W238R* mutant linked to PCH1b [37, 41-43, 54]. The *rrp4-G226D* cells exhibit a more profound growth defect than *rrp40-W195R* cells at 37°C as determined by comparing each mutant to the corresponding wild-type control (Figure 3B, 3C).

The growth defects associated with the *rrp4* mutant cells could be due to a decrease in the level of the essential Rrp4 protein. To explore this possibility, we examined the expression of Myc-tagged wild-type Rrp4 and Rrp4 variants by immunoblotting and quantitated the changes in steady-state level of Rrp4 G226D-Myc and Rrp4 G58V-Myc compared to wild-type control (Figure 3D-F). We first examined the steady-state levels of Myc-tagged Rrp4 G226D when expressed as the sole copy of the Rrp4 protein in *rrp4Δ* cells grown at either 30°C or 37°C. Immunoblotting reveals that the steady-state level of Rrp4 G226D is comparable to wild-type Rrp4 at 30°C; however, at 37°C, the level of Rrp4 G226D is decreased to ~75% of that of wild-type Rrp4 (Figure 3D). As Rrp4 G58V does not support cell viability, we could not examine the expression of this variant as the sole copy of Rrp4 in cells. Thus, we examined the expression of Myc-tagged Rrp4, Rrp4 G58V, and Rrp4 G226D in the presence of *RRP4*. Under these conditions, where an untagged copy of Rrp4 is present, the steady-state level of Rrp4 G58V-Myc is decreased to ~68% that of wild-type Rrp4 and Rrp4 G226D-Myc is decreased to ~51% that of the wild-

type Rrp4 at both 30°C and 37°C (Figure 3E). These data show that the level of Rrp4 G58V is not decreased more than Rrp4 G226D in the presence of a wild-type copy of Rrp4. Quantitation of results from these studies are shown in Figure 3F and 3G. The Rrp4 variants show a decrease in steady-state levels in the presence of wild-type Rrp4, suggesting cells can discriminate between wild-type and variant RNA exosome subunits.

The Rrp4 G226D variant can associate with the RNA exosome complex *in vivo* in the absence of competing wildtype Rrp4

The amino acid substitutions in the Rrp4 variants could decrease association of Rrp4 with the other cap and/or core subunits of the RNA exosome, as reported for human EXOSC2 G30V and EXOSC2 G198D [47]. To initially examine the association of Rrp4 variants with the RNA exosome, we performed co-immunoprecipitations using *RRP43-TAP* cells that contain the endogenous *RRP4* gene and express a C-terminally tandem affinity purification (TAP)-tagged Rrp43 core subunit from the endogenous *RRP43* locus. We co-expressed Rrp4-Myc, Rrp4 G58V-Myc or Rrp4 G226D-Myc in these *RRP43-TAP* cells. The Rrp43-TAP protein was immunoprecipitated and association of the Myc-tagged Rrp4 variants was assayed by immunoblotting (Figure 4A). Under these conditions where an endogenous, wild-type copy of *RRP4* is present, we do not detect association of Rrp4 G58V-Myc or Rrp4 G226D-Myc with Rrp43-TAP, whereas we do detect association of the wild-type Rrp4-Myc with Rrp43-TAP (Figure 4A).

To further investigate association of Rrp4 G226D with the RNA exosome complex, we performed the same co-immunoprecipitation experiment in the absence of endogenous *RRP4* using *RRP43-TAP rrp4Δ* cells that express Rrp4-Myc or Rrp4 G226D-Myc. We could not express Rrp4 G58V-Myc in these cells as Rrp4 G58V does not support viability (Figure 3A). The Rrp43-TAP protein was immunoprecipitated and association of Rrp4-Myc or Rrp4 G226D-Myc was assayed by immunoblotting (Figure 4B). Under these conditions where Rrp4 G226D-Myc is the sole copy of the essential cap subunit, we detect association with the RNA exosome complex at levels comparable to wild-type Rrp4-Myc (Figure 4B). These data suggest that Rrp4 G226D can associate with the RNA exosome complex when it

is the sole copy of the cap subunit; however, in the presence of endogenous *RRP4*, the wild-type copy of Rrp4 can outcompete pathogenic Rrp4 variants for incorporation into the RNA exosome complex.

The Rrp4 G226D variant impairs RNA exosome function

To assess the function of the RNA exosome in *rrp4-G226D* cells, we examined the steady-state level of several well-defined RNA exosome target transcripts. The RNA exosome has a critical role in ribosomal RNA (rRNA) processing, specifically processing 7S pre-rRNA into mature 5.8S rRNA [2, 9]. We analyzed the processing of 5.8S rRNA in *rrp4-G226D* cells using northern blotting. We also compared 5.8S rRNA processing in *rrp4-G226D* cells to yeast cells modeling *EXOSC3* PCH1b mutations, *rrp40-G8A* and *rrp40-W195R* [49, 51]. As shown in Figure 5A, *rrp4-G226D* cells accumulate 7S pre-rRNA, a precursor of mature 5.8S rRNA. In addition, several intermediate precursors of 5.8S rRNA, indicated by asterisks, accumulate in *rrp4-G226D* cells. Despite the accumulation of precursors, the level of mature 5.8S rRNA does not appear to differ in *rrp4-G226D* cells compared to control *RRP4* cells. The accumulation of 7S pre-rRNA and other 5.8S rRNA precursors in *rrp4-G226D* cells is greater than that detected in *rrp40-W195R* cells (Figures 5A, S3), which have documented accumulation of this rRNA precursor [51].

We next analyzed the steady-state levels of several RNA exosome target transcripts in *rrp4-G226D* cells using quantitative RT-PCR [9]. The *rrp4-G226D* cells exhibit a significant increase in the level of 3'-extended *U4* pre-snoRNA compared to *RRP4* control cells (Figure 5B). The *rrp4-G226D* cells exhibit a significant increase in the level of *U14* box C/D snoRNA, whereas they show no significant difference in the level of the *snR44* box H/ACA snoRNA (Figure 5C). We also measured the steady-state level of the telomerase RNA *TLC1*, which is processed by the RNA exosome in a pathway similar to pre-snoRNA processing [55]. The *rrp4-G226D* cells exhibit a significant increase in the steady-state level of both the mature and the extended precursor form of *TLC1* compared to *RRP4* cells (Figure 5D). These data indicate that known RNA exosome target transcripts accumulate in *rrp4-G226D* cells and suggest that Rrp4 G226D impairs the function of the RNA exosome.

The Rrp4 G226D variant causes broad transcriptomic changes

To further investigate the molecular consequences of the Rrp4 G226D substitution, we performed RNA-Seq analysis on rRNA-depleted total RNA isolated from three independent biological replicates of the *rrp4-G226D* and control *RRP4* cells as described in *Materials and Methods*. Unbiased principal component analysis (PCA) of the resulting RNA-Seq data produced two distinct clusters, indicating that the *rrp4* mutant transcriptome is distinct from the wild-type *RRP4* control (Figure 6A). This separation between the two genotypes and reproducibility amongst the RNA-Seq replicates allowed us to identify transcriptomic changes in *rrp4-G226D* mutant cells compared to the control (Figure 5B). From differential gene expression analysis, we detect 860 transcripts increased ($\geq +1.5$ Fold Change [FC], $p < 0.05$) and 802 transcripts decreased (≥ -1.5 FC, $p < 0.05$) in *rrp4-G226D* cells compared to the *RRP4* control (Figure 6B). Of the 860 transcripts increased, only a third are mRNAs (34%, 296 transcripts), with the majority being cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs), and other ncRNAs (Figure 6C). Consistent with the role the RNA exosome plays in degradation of nascent ncRNA species, the CUTs and SUTs combined make up the majority (65%) of transcripts that show a steady-state increase in *rrp4-G226D* cells (Figure 6C). Of the 802 transcripts decreased, the majority are mRNAs (90%, 719 transcripts) (Figure 6C), with the most significantly decreased transcript (≥ -4 FC) being *INO1*, an mRNA that encodes inositol-3-phosphate synthetase [56, 57], which has previously been identified as a transcript bound to the catalytic subunit of the RNA exosome [58].

Gene Ontology (GO) analysis of the differentially expressed transcripts in *rrp4-G226D* cells using YeastEnrichr [59-61] reveals that ncRNA catabolic process is the most significant biological process category for the increased transcripts (Combined score 19.56) and cytoplasmic translation is the most significant category for the decreased transcripts (Combined score 600.4) (Figure 6D). These GO analyses align with the transcripts that are altered, as two significantly decreased mRNAs (≥ -1.5 FC), *RPS3* and *RPL15A*, encode components of the ribosome, and two significantly increased mRNAs ($\geq +1.5$

FC), *NRD1* and *NAB3*, encode components of the Nrd1-Nab3-Sen1(NNS) transcription termination complex [62-65].

To validate altered gene expression detected in the RNA-Seq analysis, we measured the levels of a subset of transcripts (Figure 7). We performed this analysis on select coding and non-coding transcripts (labeled in Figure 6B). This analysis confirms that the steady-state levels of three non-coding CUTs — *CUT501*, *CUT770*, *CUT896* (Figure 7A) — and three mRNAs — *PTH4*, *NRD1*, *NAB3* (Figure 7C, 7D)— that increased in the RNA-Seq analysis are significantly increased ($p < 0.05$) in *rrp4-G226D* cells compared to *RRP4* control cells. We also validated decreased steady-state levels of several mRNAs (*RPS3*, *RPL15A*, *INO1*, *HXK2*, *TDH1*) ($p < 0.01$) in *rrp4-G226D* cells compared to control (Figure 7B, C).

To compare the molecular consequences resulting from two pathogenic missense mutations in RNA exosome cap subunits, EXOSC2/Rrp4 and EXOSC3/Rrp40 [49, 51], we expanded the RT-qPCR analysis to include *rrp40-W195R* cells. Intriguingly, we found that some altered targets in *rrp4-G226D* cells are affected in both mutants, while others are significantly affected only in the *rrp4* mutant. The steady-state levels of *CUT501*, *CUT770*, and *CUT896* are only significantly increased in *rrp4-G226D* cells and not in *rrp40-W195R* cells (Figure 7A). Steady-state levels of coding *RSP3*, *RPL15A*, and *INO1* mRNAs are significantly decreased in both *rrp4-G226D* and *rrp40-W195R* cells compared to control cells (Figure 7B). In contrast, the decrease in steady-state levels of *HXK2* and *TDH1* mRNAs is unique to the *rrp4-G226D* cells, as these coding RNAs are not affected in *rrp40-W195R* cells (Figure 7C). The *PTH4* mRNA is significantly increased in *rrp40-W195R* cells compared to *RRP40* control, as observed in *rrp4-G226D* cells; however, the magnitude of the change detected was quite different. With respect to the NNS components, the *NRD1* steady-state level changes to a similar extent in both *rrp4-G226D* and *rrp40-W195R* cells compared to control; however, the significant increase in the *NAB3* steady-state level occurs only in *rrp4-G226D* cells (Figure 7D).

The *rrp4-G226D* mutant shows genetic interactions with nuclear RNA exosome cofactors

The specificity of the RNA exosome for different RNA substrates is conferred by several interacting cofactors, which have been most extensively characterized in budding yeast [15, 22]. As depicted in Figure 8A, the exosome cofactor Rrp47 interacts with and stabilizes the exoribonuclease Rrp6, and the cofactor Mpp6 directly interacts with the nuclear RNA exosome [18, 26, 27]. To determine whether the *rrp4-G226D* variant exhibits genetic interactions with RNA exosome cofactor mutants, we deleted the non-essential, nuclear exosome cofactor genes *MPP6*, *RRP47* as well as *RRP6* in combination with *rrp4-G226D*. For comparison, we also tested whether the *rrp40-W195R* variant shows genetic interactions with this series of mutants by deleting each gene in combination with *rrp40-W195R*. We examined the growth of these double mutants relative to single mutants (*rrp4-G226D* and *rrp40-W195R*) in solid media growth assays (Figure 8). Interestingly, the *rrp4-G226D mpp6Δ*, *rrp4-G226D rrp6Δ*, and the *rrp4-G226D rrp47Δ* double mutant cells all exhibit impaired growth compared to *rrp4-G226D* and single mutants at 30°C (Figure 8A), indicating that deletion of *MPP6*, *RRP47* or *RRP6* exacerbates the growth defect of *rrp4-G226D* cells. The impaired growth of the *rrp4-G226D rrp6Δ* double mutant is particularly striking. In contrast, *rrp40-W195R mpp6Δ*, *rrp40-W195R rrp47Δ*, and *rrp40-W195R rrp6Δ* double mutant cells do not show altered growth compared to *rrp40-W195R* or single mutant cells at 30°C (Figure 8B).

The *rrp4-G226D* double mutants also exhibit enhanced growth defects relative to single mutants at 37°C. The impaired growth of the *rrp4-G226D mpp6Δ* double mutant at 37°C is particularly noteworthy as loss of *MPP6* does not alter cell growth at either 30°C or 37°C in single mutant cells or in double mutant *rrp40-W195R* cells (Figure 8A). The *rrp40-W195R rrp47Δ* and *rrp40-W195R rrp6Δ* double mutant cells exhibit impaired growth at 37°C, though not substantially worse when compared to the impaired growth of single mutants *rrp47Δ* or *rrp6Δ* at 37°C, as has been previously reported [17, 66] (Figure 8B).

Rrp4 G226D has decreased association with the essential helicase Mtr4

The nuclear exosome cofactors Mpp6 and Rrp47 and the associated exoribonuclease Rrp6 aid in recruiting the essential nuclear RNA helicase, Mtr4, to the RNA exosome [27] (Figure 9A). Rrp6 and Rrp47 form a composite site that binds to the N-terminus of Mtr4, recruiting the helicase to the RNA exosome [26]. Mpp6 also tethers Mtr4 and Mtr4-containing complexes to the complex [25]. Structural studies have also shown that human MTR4/MTREX and budding yeast Mtr4 directly interacts with the RNA exosome by binding to a conserved region of EXOSC2/Rrp4 that also facilitates EXOSC10/Rrp6-RNA exosome interaction [28, 29] (Figure 9A). Given the negative genetic interactions observed between *rrp4-G226D* and nuclear cofactor mutants, and the binding interface between EXOSC2 and human MTR4 shown in structural studies [28], we tested whether the interaction between the RNA exosome and Mtr4 is affected in *rrp4-G226D* cells.

We tested for a genetic interaction between *rrp4-G226D* and *mtr4-F7A-F10A*, a mutant allele of *MTR4* that disrupts interactions with Rrp6/Rrp47 [26] (Figure 9B). The *rrp4-G226D mtr4-F7A-F10A* double mutant cells grow similarly to *rrp4-G226D* cells at 30°C, however, the *rrp4-G226D mtr4-F7A-F10A* cells are not viable at 37°C (Figure 9B). As a comparison, we performed a similar growth assay with *rrp40-W195R mtr4-F7A-F10A* double mutant cells and found that these cells show growth similar to *rrp40-W195R* cells at 30°C and 37°C (Figure 9B). These data show that *rrp4-G226D* genetically interacts with a mutant allele of *MTR4* that disrupts interactions with Rrp6/Rrp47.

To investigate whether the physical interaction between Mtr4 and the RNA exosome is impacted by the Rrp4 G226D variant, we performed a co-immunoprecipitation with cells that express Rrp4-Myc or Rrp4 G226D-Myc as the sole copy of Rrp4 and co-express Mtr4-FLAG (Figure 9C-E). The Rrp4-Myc proteins were immunoprecipitated and association with Mtr4-FLAG was assayed by immunoblotting. Mtr4-FLAG co-immunoprecipitates with Rrp4-Myc but not with Rrp4 G226D-Myc (Figure 9C). Results from three independent experiments are quantified for the amount of co-isolated Mtr4-FLAG in Figure 9D. This difference in association of Mtr4-FLAG with wild-type Rrp4 versus Rrp4 G226D is not due to decreased protein levels or inefficient immunoprecipitation of Rrp4 G226D-Myc as quantitated in Figure

9E. Rather, these data demonstrate that Mtr4 association with the Rrp4 cap subunit is significantly disrupted by the Rrp4 G226D amino acid substitution (Figure 9C, D). Combined with the genetic data (Figure 9B), these results suggest that there is a disruption between Mtr4 and the RNA exosome complex in *rrp4-G226D* cells, thus providing a potential molecular mechanism for the impairment in RNA exosome caused by the Rrp4 G226D amino acid substitution.

2.5 Discussion

In this study, we modeled and analyzed pathogenic amino acid substitutions in the *S. cerevisiae* EXOSC2 ortholog, Rrp4. We generated *rrp4-G58V* and *rrp4-G226D* mutants, which correspond to the SHRF-linked mutations *EXOSC2-G30V* and *EXOSC2-G198D*, respectively. Analysis of the *rrp4-G58V* and *rrp4-G226D* cells reveals that these amino acid substitutions have distinct effects on RNA exosome function. The Rrp4-G58V variant is not able to function as the sole copy of the essential Rrp4 RNA exosome cap subunit as *rrp4-G58V* cells are not viable. In contrast, *rrp4-G226D* cells show a growth defect at 37°C. These *rrp4-G226D* cells show significant transcriptomic changes compared to wild-type cells, including increases in steady-state levels of known direct RNA exosome targets such as precursors of 5.8S ribosomal RNA (rRNA), *U4* small nuclear RNA (snRNA), and *TLC1* telomerase RNA [2, 9-12, 55, 67-69]. RNA-Seq analysis of *rrp4-G226D* cells show broad transcriptomic changes, with predominantly increased steady-state levels of non-coding RNA CUTs and SUTs that are usually regulated by nuclear surveillance mechanisms. The Rrp4 G226D variant can assemble into the RNA exosome, but both genetic and biochemical studies suggest interactions with key RNA exosome cofactors are impaired in *rrp4-G226D* cells. In particular, we observe decreased association of the essential helicase Mtr4 with Rrp4 G226D, suggesting a decreased interaction with this nuclear cofactor *in vivo*. Overall, these data suggest that the SHRF-linked pathogenic amino acid substitutions alter the overall function of the RNA exosome *in vivo*, resulting in defects in nuclear surveillance that may be due to impaired interaction with Mtr4. These results provide the first *in vivo* model of pathogenic amino acid substitutions that occur in EXOSC2.

In assessing the molecular and functional consequences of these *rrp4* variants *in vivo*, we first tested whether the modeled SHRF pathogenic amino acid substitutions affect protein levels and/or incorporation of the cap subunit into the RNA exosome. These data complement prior biochemical studies that employed SHRF patient cells [47]. Our biochemical assays suggest that Rrp4 G58V and Rrp4 G226D are not able to incorporate into the RNA exosome complex when a wild-type copy of the cap subunit is present (Figure 4A); however, Rrp4 G226D incorporates into the complex when no competing wild-type subunit is present (Figure 4B). These analyses suggest that the Rrp4 variants are outcompeted by a wild-type Rrp4 for incorporation into the complex, which is similar to previous studies of *rrp40-W195R* cells showing that Rrp40 variants cannot incorporate into the complex in the presence of a wild-type copy of Rrp40 and are subsequently targeted by the proteasome for degradation [49]. This reported decrease in protein half-life of unincorporated subunits into the complex could explain the decrease in steady-state level of Rrp4 G58V and Rrp4 G226D in cells expressing a wild-type *RRP4* copy (Figure 3E). The lethality observed in *rrp4-G58V* cells when Rrp4 G58V is the sole copy of the essential cap subunit (Figure 3A) could mean that Rrp4 G58V cannot associate with the RNA exosome, perhaps resulting in loss of functional complex *in vivo*. In contrast, *rrp4-G226D* cells are viable, but show temperature sensitive growth.

Two of the three SHRF patients identified thus far are homozygous for the missense mutation *EXOSC2-G30V* [40], suggesting that this *EXOSC2* variant is able to support RNA exosome function in humans. From our structural modeling, we predict similarities between both the *EXOSC2-EXOSC4* and *Rrp4-Rrp41* interface (Figure 2A, 2C) and a conserved stabilizing salt bridge between the two RNA exosome subunits that depends on the conserved Gly30 residue in *EXOSC2* and Gly58 residue in *Rrp4* (Figure S2). The *EXOSC2-EXOSC4* and *Rrp4-Rrp41* interfaces may be differentially impacted by the valine substitution in the two eukaryotic species which could account for the differences between budding yeast *rrp4-G58V* cells and homozygous *EXOSC2-G30V* patients. Previous studies also suggest that the RNA exosome plays an important role in tissue development and human embryonic stem cell

differentiation [70, 71]. The diverse clinical presentation in patients with SHRF could reflect these key developmental roles and/or different requirements in different cell types. Thus, the differential effects observed between budding yeast *rrp4-G58V* cells and *EXOSC3-G30V* patients could be indicative of differences in developmental time points or requirements between the two eukaryotes. Integrating additional disease models across other systems will be required to define how pathogenic missense mutations differentially impact RNA exosome function in a tissue-specific manner, leading to diverse disease pathologies.

RNA-Seq analysis of the *rrp4-G226D* mutant cells revealed a broad spectrum of RNA classes that are altered in these mutant cells. The majority of the significantly increased transcripts in *rrp4-G226D* cells are comprised of the non-coding RNAs, CUTs and SUTs (65% of all up transcripts with $FC \geq +1.5$). As CUTs and SUTs are stabilized in RNA exosome mutants and cross-link to the RNA exosome [12, 13, 67, 72], we suggest that the elevated CUTs and SUTs observed in *rrp4-G226D* cells are due to impaired nuclear exosome function due to the Rrp4 G226D substitution. In contrast, the overwhelming majority of the significantly decreased transcripts are mRNAs (90% of all down transcripts with $FC \geq -1.5$), with the most significantly decreased transcript being *INO1* mRNA. Previous studies have shown that *INO1* mRNA associates with the catalytic Dis3/Rrp44 subunit of the RNA exosome in budding yeast as determined by UV cross-linking and analysis of cDNA (CRAC) analysis [58]. This published dataset also reports physical interaction of Dis3/Rrp44 with *HXK2*, *TDHI*, *RPL15A* and *RPS3* mRNAs [58], other RNAs decreased in our RNA-Seq analysis. The authors of this study suggested that these mRNA targets could be rapidly turned over in the cytoplasm by the RNA exosome, but our results show a decrease in these mRNAs, rather than the increase predicted if these transcripts were rapidly degraded by the RNA exosome. Notably, many ribosomal protein gene (*RPG*) mRNAs are decreased in *rrp4-G226D* cells and GO analysis of the decreased transcripts revealed cytoplasmic translation to be the most significantly affected biological process (Figure 5D). Consistent with these data, decreases in *RPG* mRNAs have also been observed in *rrp6Δ* mutant cells [73]. Decreases of these specific mRNAs in *rrp4-*

G226D cells could reflect dysregulation of the cytoplasmic RNA exosome in *rrp4-G226D* cells, however it is not clear how these mRNA targets that physically associate with the catalytic subunit Rrp44 are decreased in *rrp4-G226D* cells. Overall, these data demonstrate that the Rrp4 *G226D* variant could alter both nuclear and cytoplasmic roles of the RNA exosome.

Many of the decreased mRNA transcripts observed in *rrp4-G226D* cells could result from indirect effects, reflecting cellular changes that occur when the function of the RNA exosome is compromised, leading to numerous downstream changes. Previous work in *Drosophila melanogaster* that employed RNAi to deplete Rrp4 identified decreased levels of several transcripts encoding autophagy proteins [47]. The authors postulated that defective autophagy could contribute to SHRF pathology [47]. In our RNA-Seq analysis of *rrp4-G226D* cells, we identified 16 autophagy transcripts that were decreased -1.5-fold ($p < 0.05$) (Figure S4), which is consistent with observations in this previous study. Further studies will be required to assess whether *rrp4-G226D* cells have impaired autophagy as well as to determine whether these mRNA transcripts are direct targets of regulation by the RNA exosome.

Both genetic and biochemical assays suggest that defects in the RNA exosome function in *rrp4-G226D* cells could be due to disrupted cofactor interactions, particularly decreased association with the essential RNA helicase Mtr4. Human structural studies have shown that MTR4/MTREX binds directly not only to MPH6/MPP6 but also to a conserved region of EXOSC2 [28]. Perturbance of the MTR4/MTREX/Mtr4-RNA exosome interaction in humans and budding yeast could also impact association of the NEXT complex or TRAMP complex with the RNA exosome, thus affecting nuclear RNA quality control of several RNA classes, including processing of telomerase RNA, and degradation of cryptic ncRNAs [34]. The increased levels of CUTs, SUTs, and precursors of *U4 snRNA* and *TLC1* observed in *rrp4-G226D* cells further lends support to suggest that the Mtr4-RNA exosome interaction is impaired by Rrp4 *G226D*. This finding is consistent with a previous study that employed a structural and biochemical approach to study the budding yeast nuclear RNA exosome and the consequences of the pathogenic EXOSC3 W238R variant linked to PCH1b in the human RNA [25]. This study showed that

substitution of Arg for Trp at position 238 (W238R) in EXOSC3, corresponding to Rrp40 W195R modeled in yeast, impaired the interaction with MPP6. Taken together, our analysis of Rrp4 variants and the previous study of Rrp40 variants suggest that pathogenic amino acid substitutions in cap subunits could impair interactions with RNA exosome cofactors, suggesting a molecular mechanism that could underlie impaired RNA exosome function *in vivo*.

This study identified several *in vivo* consequences resulting from the Rrp4 G226D amino acid substitution. Genetic and biochemical analyses show that *rrp4-G226D* cells have impaired RNA exosome function, broad transcriptomic changes and defects in RNA exosome cofactor interactions. When we compare these functional and molecular consequences in *rrp4-G226D* cells to *rrp40-W195R* cells we see some similarities and some differences. Both exosomopathy mutant models show growth defects, though *rrp4-G226D* cells show a more severe growth phenotype (Figure 3B). In addition, both exosomopathy mutant models share changes in steady state levels of some transcripts (Figure 7B), but some transcripts show statistically significant changes only in *rrp4-G226D* cells. Genetic analyses also suggest that *rrp4-G226D* and *rrp40-W195R* mutants have similar negative genetic interactions with key nuclear exosome cofactor mutants (Figure 8). However, we also detect a negative genetic interaction between *rrp4-G226D* and an *mtr4* mutant that is not observed with *rrp40-W195R*. Based on these observations, two possibilities present themselves. One possibility is that missense mutations in *RRP4* and *RRP40* have distinct functional consequences for RNA exosome activity, which would be consistent with the distinct clinical presentations in patients with these pathogenic variants. Alternatively, the *rrp4-G226D* allele may simply be a stronger allele than *rrp40-W195R*. A more extensive, comparison of additional exosomopathy mutant alleles will be required to distinguish between these two possibilities.

Utilizing the yeast genetic model system, we have begun to elucidate the functional consequences that result from the pathogenic amino acid substitutions in EXOSC2 in SHRF patients. By modeling these mutations in the corresponding *RRP4* gene, we have generated a system that can be used to understand how pathogenic amino acid substitutions impact the function of the RNA exosome. This study also adds

to the growing collection of *in vivo* RNA exosomopathy mutant models that can be compared to one another to define the *in vivo* consequences resulting from each mutation. For several RNA exosomopathies, including SHRF syndrome, the patient population is quite small, making analysis of patient tissue samples challenging if not impossible. Our findings presented here can be integrated into the body of work describing the SHRF *EXOSC2* mutations, further expanding our understanding of the unique disease pathology. Future comparative *in vivo* analysis of exosomopathy models will allow for deeper understanding of how diverse clinical symptoms are linked to changes in a single molecular machine. Furthermore, this type of *in vivo* comparison may shed light on the basic biology of the RNA exosome, as many questions still exist regarding RNA substrate targeting and regulation of this essential molecular machine. This study not only provides the first *in vivo* study that models *EXOSC2* mutations identified in SHRF patients, but also provides a platform for the first direct comparison of the consequences of pathogenic missense mutations in genes encoding cap subunits of the RNA exosome.

2.6 Chapter II Acknowledgements

We thank members of the Corbett and van Hoof laboratories for critical discussions and input. We also thank Dr. Graeme Conn for his intellectual contributions. This work was supported by National Institutes of Health (NIH) R01 grants (GM058728) to A.H.C. and A.v.H. and (GM099790) to A.v.H. M.C.S. was supported by a National Institute of General Medical Sciences (NIGMS) F31 grant (GM134649-01). L.E. was supported by the NIH- funded Emory Initiative for Maximizing Student Development (R25 GM099644). S.E.S. was supported by the National Science Foundation (NSF) Graduate Research Fellowship (GRFP 1937971). M.A.B is supported by the NIH-funded Intramural Research Program of the National Cancer Institute. Lastly, we would also like to thank the RNA society and Saccharomyces Genomic Database (SGD) [90] for providing community resources and support for scientific discovery.

2.7 Chapter II Figures

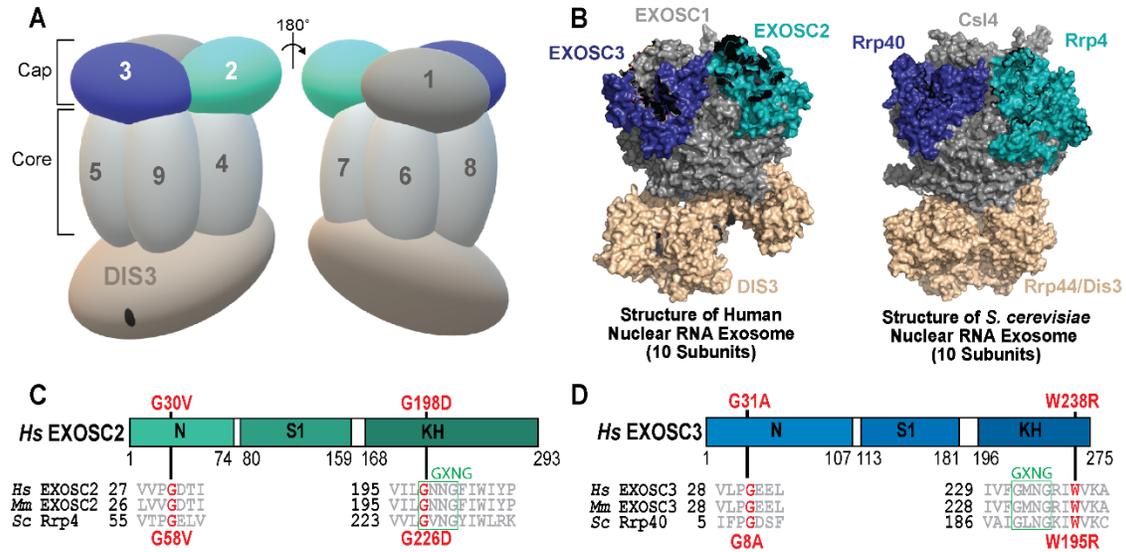


Figure 1. Overview of pathogenic amino acid substitutions in the human cap subunit EXOSC2 of the RNA exosome.

(A) The RNA exosome is an evolutionary conserved ribonuclease complex composed of nine structural subunits (EXOSC1-9) and one catalytic subunit (DIS3) that form a “cap” and “core” ring-like structure. The 3-subunit cap at the top of the complex is composed of EXOSC1/Csl4 (Human/*S. cerevisiae*), EXOSC2/Rrp4, and EXOSC3/Rrp40. The 6-subunit core is composed of EXOSC4/Rrp41, EXOSC5/Rrp46, EXOSC6/Mtr3, EXOSC7/Rrp42, EXOSC8/Rrp43, and EXOSC9/Rrp45. The DIS3/Dis3/Rrp44 catalytic subunit is located at the bottom. Missense mutations in the gene encoding the EXOSC2 cap subunit (teal blue, labeled 2,) are linked to a novel syndrome termed SHRF (short stature, hearing loss, retinitis pigmentosa and distinctive facies) [40]. In contrast, missense mutations in the gene encoding the EXOSC3 cap subunit (dark blue, labeled 3) cause PCH1b (pontocerebellar hypoplasia type 1b) [37, 41-43, 54]. (B) The structure and organization of the RNA exosome is highly conserved across eukaryotes. A structural model of the human nuclear RNA exosome (left) [PDB 6D6R] [28] and the *S.*

cerevisiae nuclear RNA exosome (right) [PDB 6FSZ] [29] are depicted with the cap subunits EXOSC1/Csl4 (Human/*S. cerevisiae*), EXOSC2/Rrp4, EXOSC3/Rrp40, and catalytic subunit DIS3/Dis3/Rrp44 labeled. (C,D) Domain structures are shown for (C) EXOSC2/Rrp4 and (D) EXOSC3/Rrp40. Each of these cap subunits is composed of three different domains: an N-terminal domain, a putative RNA binding S1 domain, and a C-terminal putative RNA binding KH (K homology) domain. The “GxNG” motif identified in the KH domain of both cap subunits is boxed in green. The position of the disease-linked amino acid substitutions in human EXOSC2 and EXOSC3 are depicted above the domain structures in red. Sequence alignments of EXOSC2/Rrp4 and EXOSC3/Rrp40 orthologs from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*) and *S. cerevisiae* (*Sc*) below the domain structures show the highly conserved residues altered in disease in red and the conserved sequences flanking these residues in gray. The amino acid substitutions in *S. cerevisiae* Rrp4 generated in this study and those in *S. cerevisiae* Rrp40, described previously [49, 51], that correspond to the disease-linked amino acid substitutions in human EXOSC2 and EXOSC3 are shown below the structures in red.

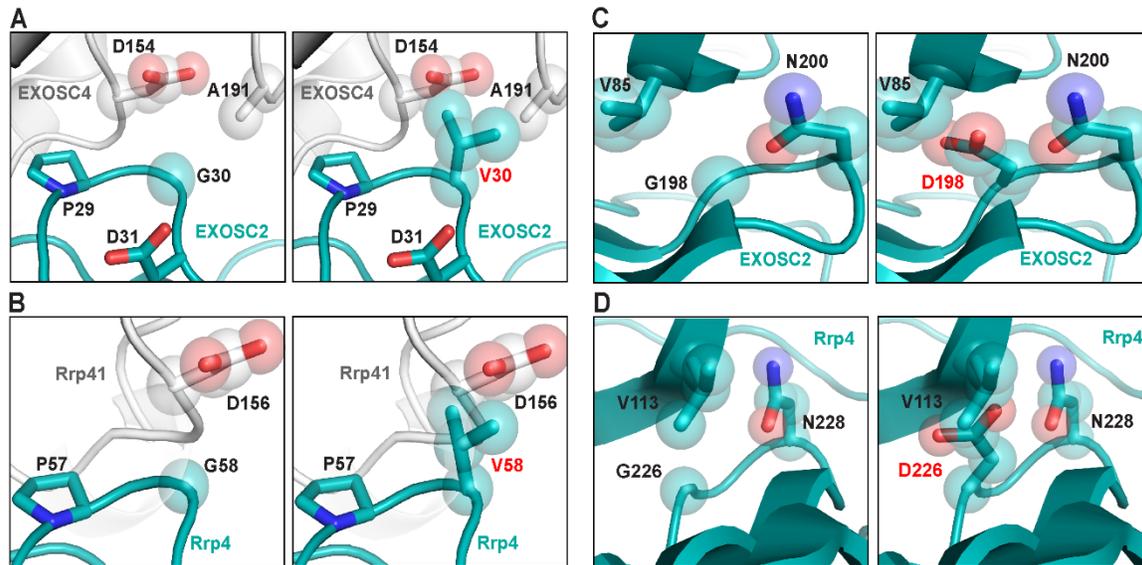


Figure 2. Modeling pathogenic amino acid substitutions in Human EXOSC2 and *S. cerevisiae* Rrp4.

(A) Structural modeling of the human EXOSC2 p.Gly30Val (G30V) amino acid substitution identified in patients with SHRF syndrome in the human RNA exosome. Zoomed-in representations of the interface between EXOSC2 (teal blue) and EXOSC4 (light gray) modeling the native EXOSC2 Gly30 (G30) residue (left) or the pathogenic EXOSC2 Val30 (V30) residue (right) are depicted. The EXOSC2 Gly30 residue is located in the N-terminal domain of EXOSC2, near the interface of EXOSC2 with the core subunit, EXOSC4. (B) Structural modeling of budding yeast Rrp4 Gly58Val (G58V) amino acid change, corresponding to EXOSC2 p.Gly30Val amino acid change, in the budding yeast exosome. Zoomed-in representations of the interface between Rrp4 (teal blue) and the budding yeast EXOSC4 ortholog, Rrp41 (light gray), modeling the native Rrp4 Gly58 (G58) residue (left) or the modeled pathogenic Rrp4 Val58 (V58) residue (right) are shown. The Rrp4 Gly58 residue is conserved between human and yeast and, similar to EXOSC2 Gly30, is located in the N-terminal domain of Rrp4, near the interface of Rrp4 with the core subunit, Rrp41. (C) Structural modeling of the EXOSC2 p.Gly198Asp (G198D) amino acid

substitution identified in patients with SHRF syndrome in the human RNA exosome. Zoomed-in representations of EXOSC2 modeling the native EXOSC2 Gly198 (G198) residue (left) or the pathogenic EXOSC2 Asp198 (D198) residue (right) are shown. The EXOSC2 Gly198 residue is located in the KH-domain of EXOSC2 within a dense region of the protein, surrounded by four β -sheets. (D) Structural modeling of the budding yeast Rrp4 Gly226Asp (G226D) amino acid change, corresponding to the EXOSC2 p.Gly198Asp amino acid change, in the budding yeast RNA exosome. Zoomed-in representations of Rrp4 modeling the native Rrp4 Gly226 (G226) residue (left) or the modeled pathogenic Rrp4 Asp226 (D226) residue (right) are shown. The Rrp4 Gly226 residue, which is conserved between human and yeast, is located in the KH-domain of Rrp4 within a dense region of the protein, surrounded by four β -sheets. Structural modeling in (A) and (C) was performed with the human RNA exosome structure (PDB 6D6R) [28] and in (B) and (D) with the yeast RNA exosome structure (PDB 6FSZ) [29] using PyMOL [75].

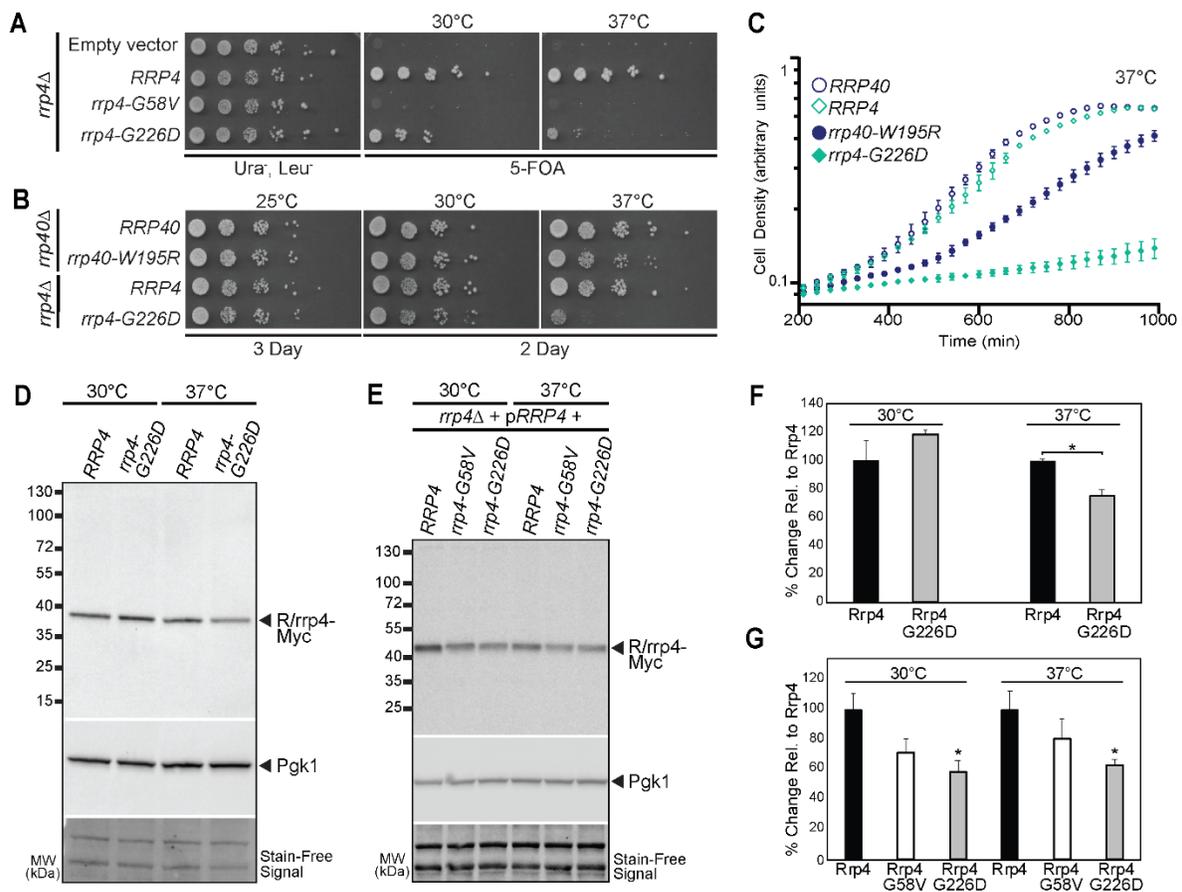


Figure 3. *S. cerevisiae* Rrp4 variants that model EXOSC2 variants identified in patients show impaired function.

S. cerevisiae cells expressing Rrp4 variants that model pathogenic amino acid changes found in EXOSC2 were generated as described in *Materials and Methods*. (A) Although cells growth is comparable for all mutants that contain a wild-type *RRP4* maintenance plasmid (Ura⁻ Leu⁻), *rrp4-G58V* mutant cells are not viable on plates containing 5-FOA where the maintenance plasmid is not present. The *rrp4-G226D* cells show temperature sensitive growth on 5-FOA relative to control *RRP4* cells. The cells were grown at the indicated temperatures. (B, C) The *rrp4-G226D* cells exhibit profoundly impaired growth compared to control *RRP4* cells at 37°C as assessed by (B) serial dilution growth assay on plates or (C) growth in liquid media. (B) The *rrp4Δ* cells expressing only *RRP4* or *rrp4-G226D* and *rrp40Δ* cells expressing only

RRP40 or *rrp40-W195R* were serially diluted, spotted onto solid media grown at the indicated temperatures or (C) grown in liquid media at 37°C with optical density measurement used to assess cell density over time. The growth of *rrp40-W195R* cells, previously reported to be moderately impaired at 37°C [49, 51], was included as a comparative control. (D) The steady-state level of the Rrp4 G226D protein variant is modestly decreased at 37°C. Lysates of *rrp4Δ* cells solely expressing Myc-tagged wild-type Rrp4 or *rrp4-G226D* grown at 30°C or 37°C were analyzed by immunoblotting with an anti-Myc antibody to detect Rrp4-Myc and an anti-Pgk1 antibody to detect 3-phosphoglycerate kinase (Pgk1) as a loading control. (E) The Rrp4-G58V protein variant is expressed and the steady-state level of the Rrp4 G226D protein variant is decreased in cells co-expressing wild-type Rrp4. Lysates of *rrp4Δ* cells co-expressing untagged wild-type Rrp4 and Myc-tagged wild-type Rrp4, Rrp4 G58V, or Rrp4 G226D grown at 30°C were analyzed by immunoblotting with an anti-Myc antibody to detect Rrp4-Myc and anti-Pgk1 antibody to detect 3-phosphoglycerate kinase (Pgk1) as loading control. (F) Quantitation of the percentage of Rrp4 or Rrp4 G226D protein detected in lysates of *rrp4Δ* cells solely expressing Myc-tagged Rrp4 or Rrp4 G226D grown at 30°C or 37°C. Graph shows the mean percentage of Rrp4-Myc protein from three independent experiments (n=3). Error bars represent standard error of mean. Statistical significance is denoted by asterisk (**p*-value ≤ 0.05). (G) Quantitation of the percentage of Rrp4, Rrp-G58V and Rrp4-G226D protein detected in lysates of *rrp4Δ* cells expressing Myc-tagged Rrp4 or Rrp4 variants grown at 30°C or 37°C. Graph shows the mean percentage of Rrp4-Myc from three independent experiments (n=3). Error bars represent standard error of mean. Statistical significance is denoted by asterisk (**p*-value ≤ 0.05). Quantitation of immunoblots in (F) and (G) was performed as described in *Materials and Methods*.

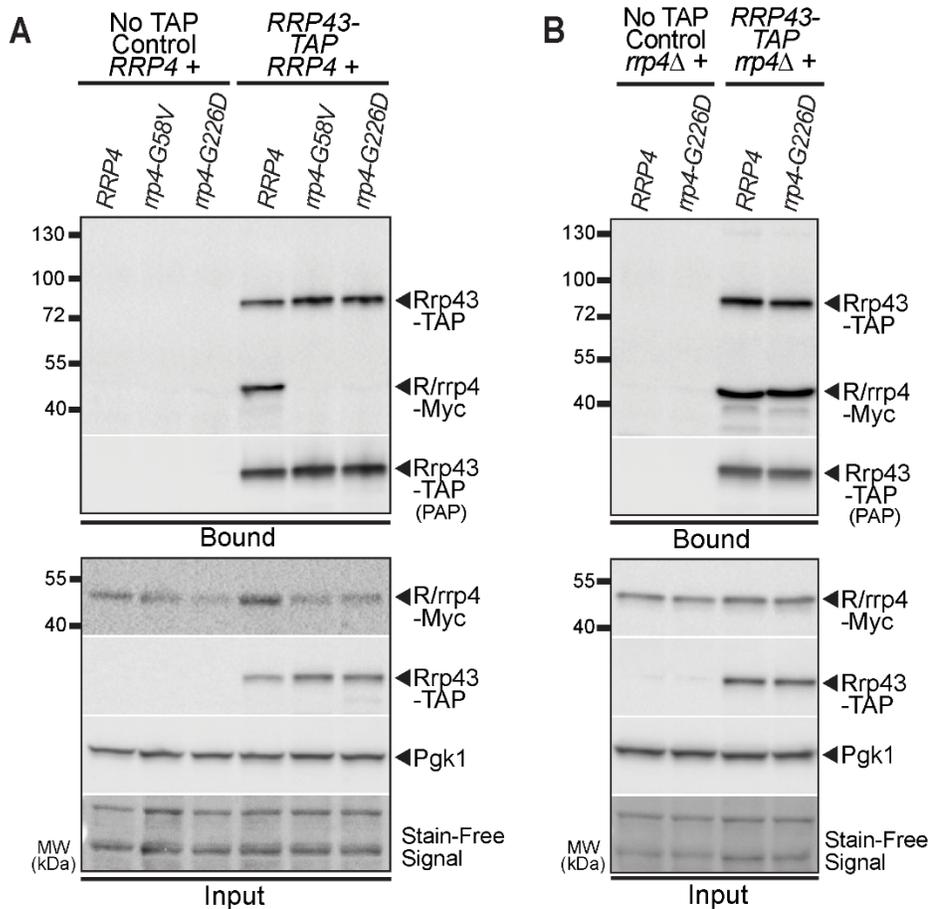
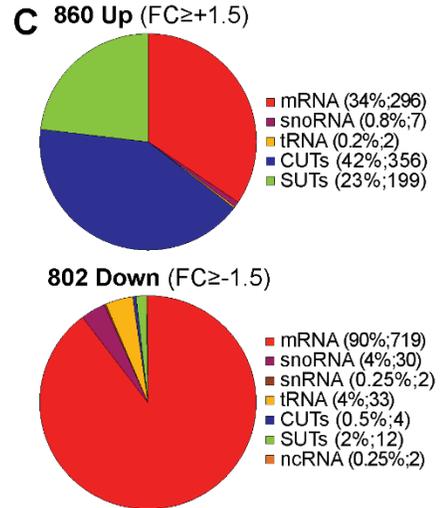
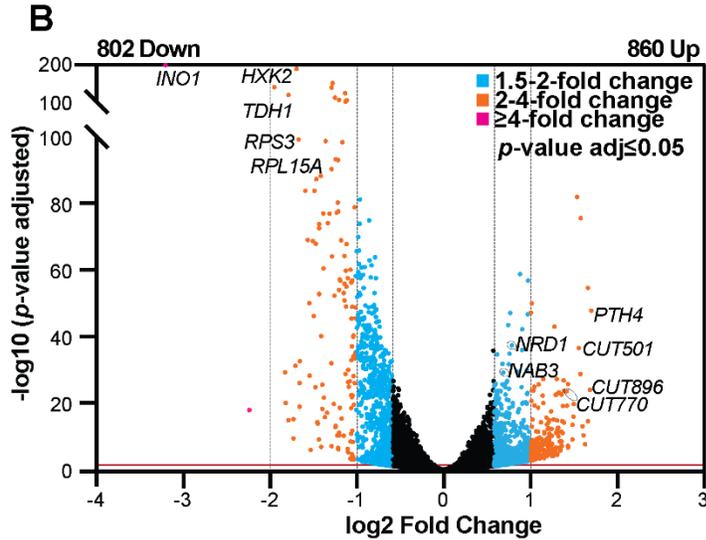
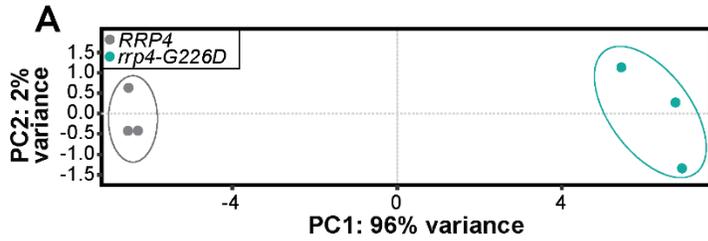


Figure 4. The Rrp4 G58V and Rrp4 G226D variants do not associate with the RNA exosome complex in the presence of wild-type Rrp4, but the Rrp4 G226D variant can associate with the RNA exosome complex.

(A) The Rrp4 G226D and Rrp4 G58V variants do not associate with the RNA exosome core subunit Rrp43 in the presence of a wild-type copy of Rrp4. TAP-tagged Rrp43 was immunoprecipitated from *RRP43-TAP* cells expressing Myc-tagged Rrp4, Rrp4 G58V, or Rrp4 G226D in the presence of wild-type Rrp4 grown at 30°C using IgG Sepharose beads and analyzed by immunoblotting. As a control,

immunoprecipitations were also performed from untagged *RRP43* cells (No TAP Control) expressing Myc-tagged Rrp4 proteins. The bound/input level of Rrp4-Myc was detected with an anti-Myc antibody and bound/input level of Rrp43-TAP was detected with a peroxidase anti-peroxidase (PAP) antibody. Bound Rrp43-TAP was also detected by the anti-Myc antibody as the Protein A moiety of the TAP tag binds to antibody. The input level of 3-phosphoglycerate kinase (Pgk1) was detected with an anti-Pgk1 antibody as a loading control. The Stain-Free signal from input protein is also included as a loading control. (B) The Rrp4 G226D variant associates with the RNA exosome core subunit Rrp43 at a level similar to wild-type Rrp4 when the variant is the sole form of Rrp4. TAP-tagged Rrp43 was immunoprecipitated from *RRP43-TAP rrp4Δ* cells expressing either Myc-tagged wild-type Rrp4 or Rrp4 G226D, which were grown at 30°C using IgG Sepharose beads and analyzed by immunoblotting. As a control, immunoprecipitations were also performed from untagged *RRP43 rrp4Δ* cells (No TAP Control) expressing Myc-tagged Rrp4 proteins. The bound/input level of Rrp4-Myc was assessed with an anti-Myc antibody and bound/input level of Rrp43-TAP was detected with a peroxidase anti-peroxidase (PAP) antibody. Bound Rrp43-TAP was also detected by the anti-Myc antibody as the Protein A moiety of the TAP tag binds to antibody. The input level of 3-phosphoglycerate kinase (Pgk1) was detected with an anti-Pgk1 antibody as a loading control. The immunoblots are representative of triplicate experiments; co-immunoprecipitations were performed as described in *Materials and Method*.

cells grown at 37°C was analyzed by northern blotting with an 5.8S-ITS2 probe to detect 7S pre-rRNA. Mature 5.8S rRNA and 5S rRNA was detected by methylene blue staining as a loading control. The 7S pre-rRNA is normally processed to mature 5.8S rRNA by 3'-5' decay of the internal transcribed spacer 2 (ITS2) via the nuclear RNA exosome [2, 9]. All lanes are imaged from the same northern blot with a gap in the loading indicated by the white line. The simplified schematics to the right illustrate the processing steps of 7S rRNA precursor following endonucleolytic cleavage from the larger 27S precursor (indicated by white triangles). (B) The *rrp4-G226D* cells show an elevated steady-state level of 3'-extended pre-*U4* snRNA relative to *RRP4* cells at 37°C. (C) The *rrp4-G226D* cells exhibit an increased steady-state level of *U14* (*snR128*) snoRNA but not *snR44* snoRNA relative to *RRP4* cells at 37°C. (D) The *rrp4-G226D* cells show an elevated steady-state level of mature and extended precursor *TLC1* telomerase component ncRNA relative to *RRP4* cells at 37°C. In (B-D), total RNA was isolated from cells grown at 37°C and transcript levels were measured by RT-qPCR using gene specific primers, normalized relative to *RRP4*, and graphed as described in *Materials and Methods*. Gene specific primer sequences are summarized in Table S2 and their location within the transcript are graphically represented by the cartoons above each bar graph. Within the cartoon transcript, the box represents the body of the mature transcript. Error bars represent standard error of the mean from three biological replicates. Statistical significance of the RNA levels in *rrp4-G226D* cells relative to *RRP4* cells is denoted by an asterisk (**p*-value ≤ 0.05; ***p*-value ≤ 0.01).



D

Up Genes: GO Biological Process	Combined Scores
antisense RNA transcript catabolic process (GO:0071041)	19.56
ncRNA catabolic process (GO:0034661)	16.39
endonucleolytic cleavage in 5'-ETS of tricistronic rRNA transcript (GO:0000480)	15.33
snoRNA 3'-end processing (GO:0031126)	13.99
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA (GO:0000472)	13.74
oligopeptide transport (GO:0006857)	13.22
anaphase-promoting complex-dependent catabolic process (GO:0031145)	13.19
rRNA 5'-end processing (GO:0000967)	13.05
endonucleolytic cleavage of tricistronic rRNA transcript (GO:0000479)	12.85
sister chromatid segregation (GO:0000819)	12.58

Down Genes: GO Biological Process	Combined Scores
cytoplasmic translation (GO:0002181)	600.4
translation (GO:0006412)	510.49
ribosome assembly (GO:0042255)	119.63
nicotinamide nucleotide metabolic process (GO:0046496)	84.86
peptide biosynthetic process (GO:0043043)	80.11
cellular macromolecule biosynthetic process (GO:0034645)	79.46
ribosomal small subunit assembly (GO:0000028)	78.84
pyruvate metabolic process (GO:0006090)	77.66
ribonucleoprotein complex assembly (GO:0022618)	76.01
glycolytic process (GO:0006096)	73.95

Figure 6. RNA-Seq analysis of *rrp4-G226D* cells reveal distinct transcriptomic changes compared to *RRP4* cells.

(A) Principal component analysis (PCA) of RNA-Seq data collected from triplicate *RRP4* and *rrp4-G226D* cell samples shows that the gene expression patterns from independent *rrp4-G226D* samples are similar and thus cluster together, but are distinct from *RRP4* samples, which also cluster together. (B) A volcano plot of the differentially expressed transcripts in *rrp4-G226D* cells compared to *RRP4* cells shows that 860 transcripts are significantly Up and 802 transcripts are Down by 1.5-fold or more in *rrp4-G226D* cells. Statistically significant fold changes in transcript levels (Down or Up) in *rrp4-G226D* cells relative to *RRP4* cells are color coded (1.5-2 FC (blue); 2-4 FC (orange); ≥ 4 FC (purple); p -value adjusted ≤ 0.05). Transcripts that were subsequently validated by RT-qPCR are labeled. (C) Pie charts of the percentages of different RNA classes within the 860 Up and 802 Down transcripts in *rrp4-G226D* cells reveal that increased transcripts are predominantly ncRNAs (CUTs; SUTs) and decreased transcripts are predominantly mRNAs. The RNA classes identified include messenger RNA (mRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), transfer RNA (tRNA), cryptic unstable transcripts (CUTs; small, non-coding RNA), stable unannotated transcripts (SUTs; small, non-coding RNA) and other non-coding RNA (ncRNA; e.g. *TLC1*), (D) Gene ontology (GO) analysis for biological process in the Up and Down transcripts in *rrp4-G226D* cells reveals that ncRNA processing is significantly represented in the Up transcripts and translation is significantly represented in the Down transcripts. GO analysis was performed on coding (mRNA) and non-coding RNA (tRNAs, snoRNAs, and snRNAs) classes using the YeastEnrichr web server [59-61]. Gray bars represent the statistical significance of the biological process categories computed using combined score listed (log of the p -value from the Fisher exact test multiplied by the z-score of the deviation from the expected rank).

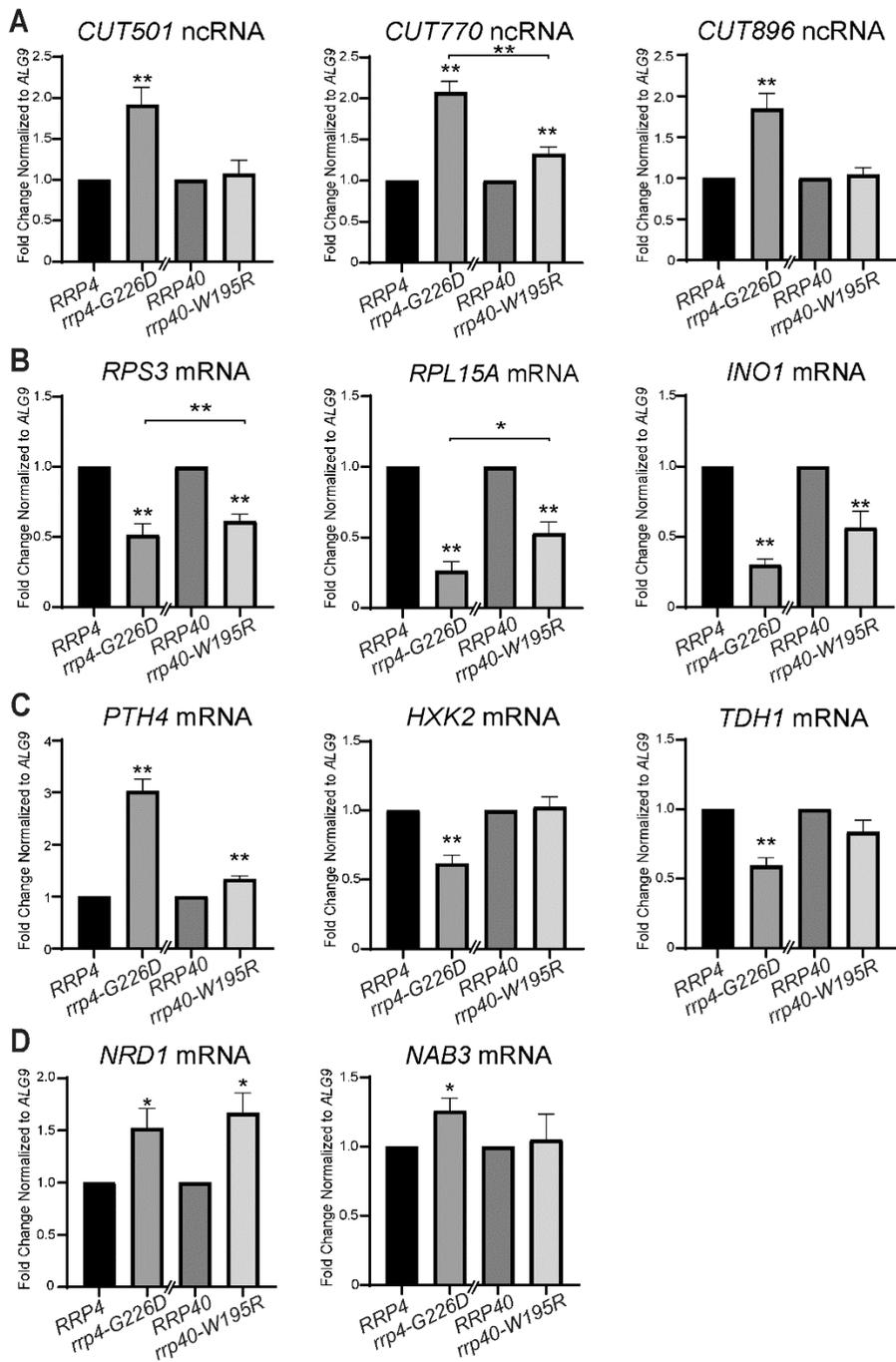


Figure 7. Validation of the differentially expressed transcripts identified in the RNA-Seq confirms that the levels of key mRNAs and CUTs are significantly altered in *rrp4-G226D* cells and reveals that some of these transcripts are not changed in *rrp40-W195R* cells.

(A) The steady-state levels of non-coding, cryptic unstable transcripts, *CUT501*, *CUT770*, and *CUT896*, are significantly increased in *rrp4-G226D* cells compared to control. The *CUT770* level is also increased but the *CUT501* and *CUT896* levels are not altered in *rrp40-W195R* cells. (B) The steady-state levels of ribosomal protein gene mRNAs, *RPS3* and *RPL15A*, and inositol-3-phosphate synthase mRNA, *INO1*, are significantly decreased in *rrp4-G226D* and *rrp40-W195R* cells relative to control *RRP4/40* cells. (C) The steady-state level of peptidyl tRNA hydrolase 4 mRNA, *PTH4*, is significantly increased in *rrp4-G226D* and *rrp40-W195R* cells relative to controls, whereas the levels of hexokinase isoenzyme 2 mRNA, *HXK2*, and glyceraldehyde-3-phosphate dehydrogenase isozyme 1 mRNA, *TDHI*, are significantly decreased in *rrp4-G226D* compared to control. The *HXK2* and *TDHI* levels are not altered in *rrp40-W195R* cells. (D) The steady-state levels of RNA exosome/termination cofactor mRNAs, *NRD1* and *NAB3*, are significantly increased in *rrp4-G226D* cells compared to controls. The *NRD1* level is increased but the *NAB3* level is not altered in *rrp40-W195R* cells. In (A-D), total RNA was isolated from cells grown at 37°C and transcript levels were measured by RT-qPCR using gene specific primers (Table S2), normalized relative to *RRP4/40*, and graphed as described in *Materials and Methods*. Error bars represent standard error of the mean from three biological replicates. Statistical significance of the RNA levels in *rrp4-G226D* and *rrp40-W195R* cells relative to control *RRP4/40* cells or between *rrp4/40* mutants is denoted by asterisks (**p*-value ≤ 0.05; ***p*-value ≤ 0.01).

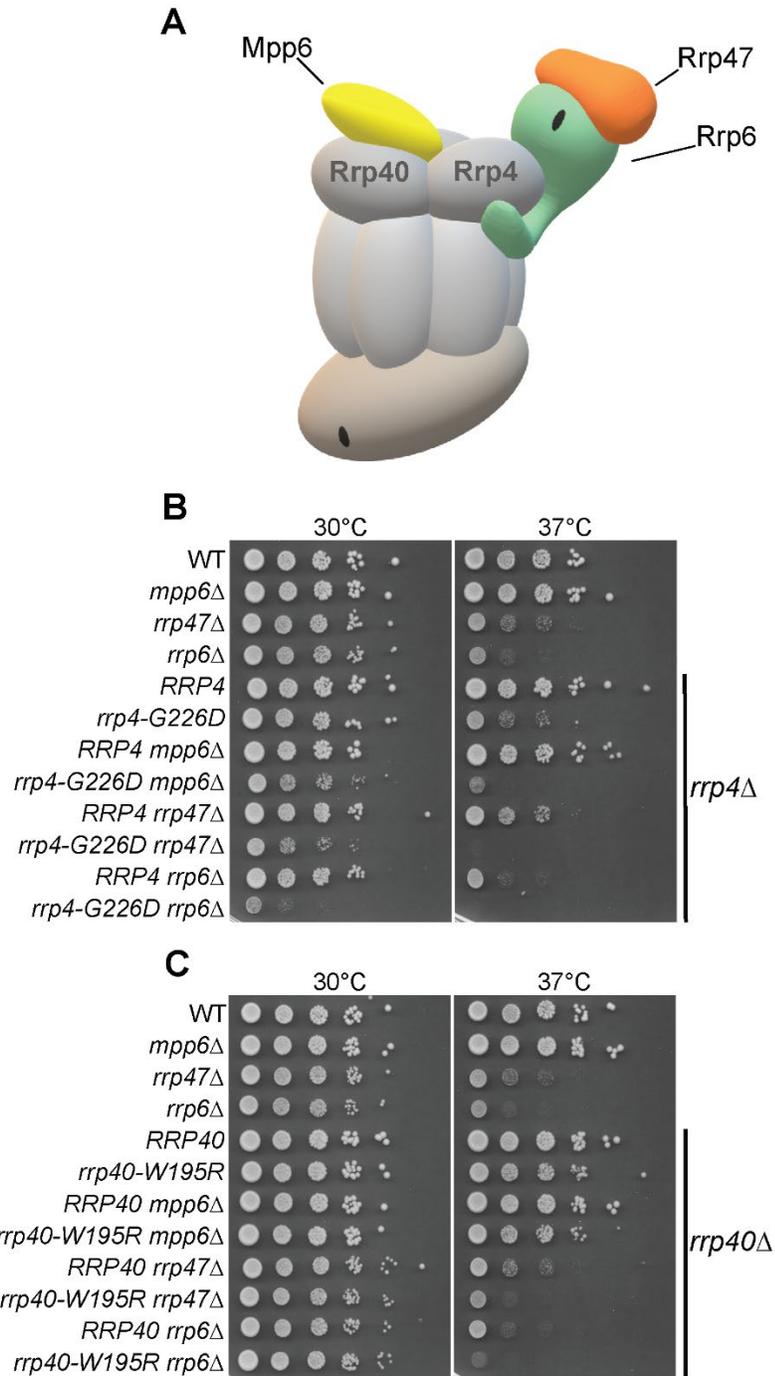


Figure 8. The *rrp4-G226D* mutant exhibits distinct negative genetic interactions with RNA exosome cofactor mutants that are not shared by the *rrp40-W195R* mutant.

(A) Cartoon schematic of the budding yeast nuclear RNA exosome interacting with nuclear cofactors, Mpp6 and Rrp47, and the exoribonuclease Rrp6 [29]. (B) Double mutant cells containing *rrp4-G226D* and *mpp6Δ*, *rrp47Δ*, or *rrp6Δ* show impaired growth compared to single mutants at 30°C and 37°C. The double mutant cells (*rrp4Δ* with *mpp6Δ*, *rrp47Δ*, or *rrp6Δ*) containing control *RRP4* or *rrp4-G226D* plasmid were serially diluted, spotted onto solid media, and grown at the indicated temperatures for 3 days. (C) Double mutant cells containing *rrp40-W195R* and *mpp6Δ* do not exhibit a change in growth compared to single mutants, whereas double mutant cells containing *rrp40-W195R* and *rrp47Δ* or *rrp6Δ* show impaired growth compared to single mutants at 37°C. The double mutant cells (*rrp40Δ* with *mpp6Δ*, *rrp47Δ*, or *rrp6Δ*) containing control *RRP40* or *rrp40-W195R* plasmid were serially diluted, spotted onto solid media, and grown at indicated temperatures for 3 days.

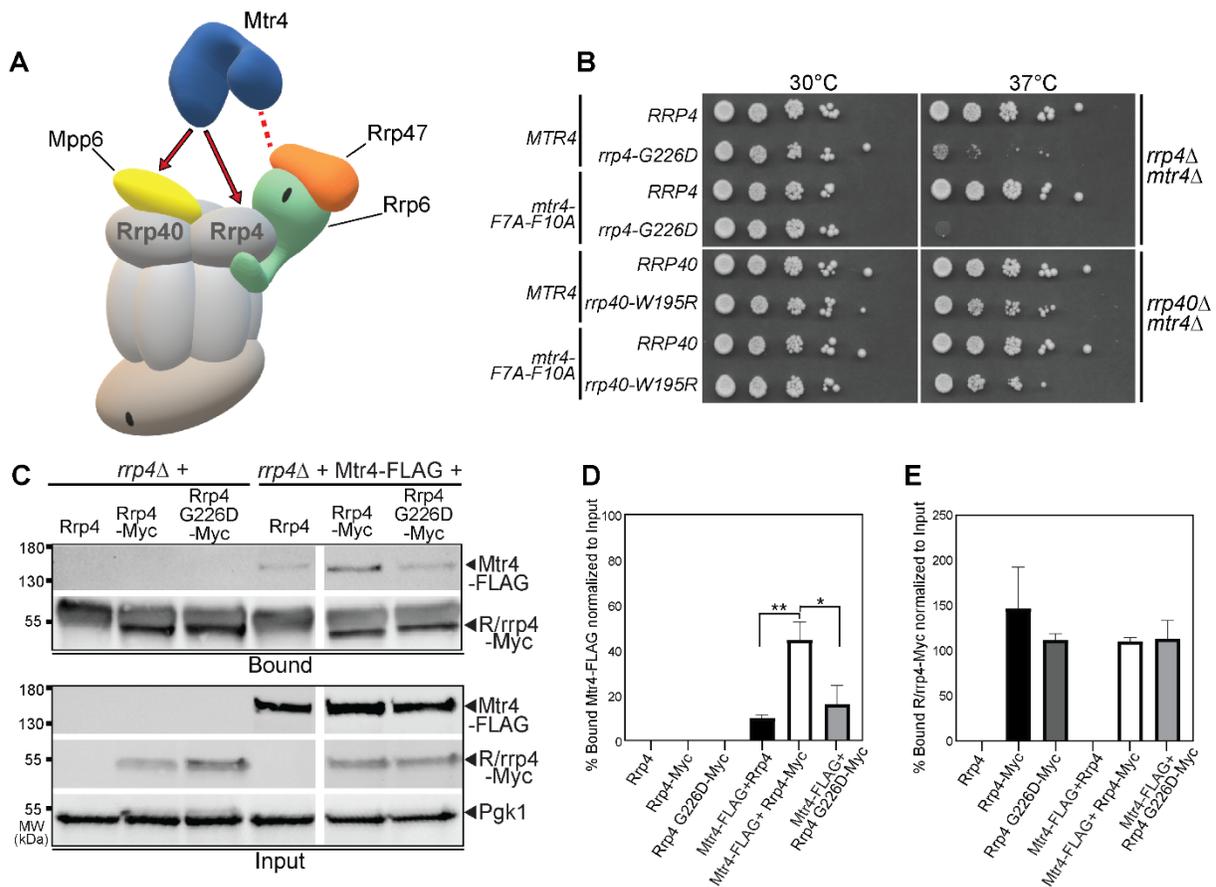


Figure 9. The *rrp4-G226D* mutant shows genetic interaction with an *mtr4* mutant that is impaired for interaction with Rrp6/Rrp47 and Rrp4 G226D impairs interaction with Mtr4.

(A) Cartoon of the budding yeast nuclear RNA exosome depicting the molecular interactions that the essential RNA helicase, Mtr4, makes with the RNA exosome and exosome cofactors [25, 26, 29]. The association of the N-terminus of Mtr4 with the RNA exosome is facilitated by interactions with nuclear exosome cofactors, Rrp6/Rrp47 (denoted by the dashed red line). The association of Mtr4 with the RNA exosome is also facilitated by interactions with nuclear exosome cofactor, Mpp6, which is associated with the Rrp40 exosome subunit, and the Rrp4 exosome subunit (denoted by the solid red arrows). (B) Double mutant cells containing *rrp4-G226D* and *mtr4-F7A-F10A*, an *mtr4* mutant impaired for interaction with Rrp6/Rrp47, show lethality compared to the impaired growth of the single mutant *rrp4-G226D* at 37°C.

In contrast, double mutant cells containing *rrp40-W195R* and *mtr4-F7A-F10A* show impaired growth at 37°C that is similar to the single mutant *rrp40-W195R*, which has been described previously [49, 51]. The *rrp4Δ mtr4Δ* double mutant cells containing *RRP4* or *rrp4-G226D* plasmid and *rrp40Δ mtr4Δ* double mutant cells containing *RRP40* or *rrp40-W195R* plasmid that also harbor *MTR4* or *mtr4-F7A-F10A* plasmid were serially diluted, spotted onto solid media, and grown at the indicated temperatures for 3 days. (C) The Rrp4 G226D variant shows decreased association with Mtr4. Myc-tagged Rrp4 and Rrp4 G226D protein was immunoprecipitated from *rrp4Δ* cells co-expressing Rrp4-Myc and FLAG-tagged Mtr4 grown at 30°C using anti-Myc beads and amount of bound Mtr4-FLAG protein was detected by immunoblotting. The bound/input level of Mtr4-FLAG was detected with an anti-FLAG antibody and the bound/input level of Rrp4-Myc was detected with an anti-Myc antibody. The input level of 3-phosphoglycerate kinase (Pgk1) was detected as a loading control. (D) Quantitation of the percentage of bound Mtr4-FLAG co-immunoprecipitated with Rrp4-Myc and Rrp4 G226D-Myc. Graph shows the mean percentage of bound Mtr4-FLAG from three independent experiments (n=3). Error bars represent standard error of the mean. Statistical significance is denoted by asterisks (**p*-value ≤ 0.05; ***p*-value ≤ 0.01). (E) Quantitation of percentage of bound Rrp4-Myc and Rrp4 G226D-Myc immunoprecipitated. Error bars represent standard error of the mean. The co-immunoprecipitations were performed and quantitated as described in *Materials and Methods*.

2.8 Chapter II Supplementary Materials

HeXOSC2	-----MAMEMRLPVARK--PL-SERLGR--DTKXHLVVP-----	-----TGFMRGHTYMGEEK-----	-----LIASVAGSERVVKL-----	-----ICVKAL
MmRrp4	-----MAEMRLPKARK--PL-SERLGR--DSKXHLVVP-----	-----TGFMRGHTYMGEEK-----	-----LIASVAGSERVVKL-----	-----ICVKAL
DrRrp4	-----MAEMRLPAVKH--HV-SLKTGQSSREDKHLVPGDITSD-----	-----TGFMRGHTYMGEEK-----	-----LIASVAGSERVVKL-----	-----ICVKPL
DmRrp4	-----MSTNAIDLALD--RVDWRDLAAQTEEPVYTPGSEVLMPE-----	-----AGFMRGHTYFEDEN-----	-----IKSVAGVQKRVNKL-----	-----ISVRPL
CeRrp4	-----MSFEVTGPP-----LVSFWMLMSSEARNIDEIKTIVPGHVSQDAP-----	-----QQFMRGHTYVRDGE-----	-----IVSSLSGVVQQLNRL-----	-----LMVKTI
ScRrp4	MSEVITITKRNQAFQNSNLSYNTTGIS--DDENDEEDIY--MHDVNSAKSESQSDIVTGGVLTDD-----	-----PIWMRGHTYFLDNM-----	-----TVSSVAGTSRVNRL-----	-----LSVIFL
SpRrp4	-----MVTILKPEEYFVSSEADIVNDVSMTEDEMEDEMEDEQGLVDGDEVLBEFDKSIHQNLVTPGQLVTD-----	-----PQFMRGHTYFEDGG-----	-----IYASVAGSQRVNKL-----	-----ISVKPL
CnRrp4	-----MFSVKAV-----APSIQAFTIIVPSTHRK--HQASLDAMDDEDEFGGAGPSKRSVSGEVISS-----	-----KEYMRGHTYVEESN-----	-----VVSVAGTIERVVKL-----	-----ISVRPL
AtRrp4	-----MVMRKLQALPISQTC--KVFERRAIE--RLQSLSS--SANSASDVIVTDSIPVNH-----	-----DAFLKRGHTSEVDGE-----	-----L LATVCGVVERVVKL-----	-----VYVRFL
OsRrp4	-----MRDLQLSLNQTQ--RVLEEARLH--ELQTVAIP--AAVAVADTIPVND-----	-----DMILKCHTSDQDGE-----	-----VVALTCGVVERVVKL-----	-----VYVRFL
HeXOSC3	-----MAEVLASGPEVAGCRRARVHVLKQVVLPEGLLPEQDEAGPGGAVERPFLSINARACSRVRCVCGPLRRCGDR-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----
MmRrp4	-----MAEVLASGPEVAGCRRARVHVLKQVVLPEGLLPEQDEAGPGGAVERPFLSINARACSRVRCVCGPLRRCGDR-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----	-----LLVTKCGRLRHKEPGSSGGGVVWDS-----
DrRrp4	-----MDSVHTLSLRLIGDVLVPEGLLPSFPPE-----AGDANPKADRLCCGGLRRSGAE-----	-----TRVCRAGVLKHKQPN-----	-----TRVCRAGVLKHKQPN-----	-----MYWVNC
DmRrp4	-----MSATSTIVMPERIAA-----IEELAKSKRVILGGLRRLDDT-----	-----VVASKAGPLRHKPEG-----	-----VVASKAGPLRHKPEG-----	-----TFWVND
CeRrp4	-----MATVILPGDVIPESS-----SDSSIIGGIVSRGQT-----	-----RIATQPGAFH-----	-----RIATQPGAFH-----	-----NDDQKVLNV
ScRrp4	-----MSTIIPGDFVDP-----MSTIIPGDFVDP-----	-----TRPVNTGLVHVSAGK--KSGVQTAYIDY-----	-----TRPVNTGLVHVSAGK--KSGVQTAYIDY-----	-----TRPVNTGLVHVSAGK--KSGVQTAYIDY-----
SpRrp4	-----MAEVEKGLLSYVFGERRIPDSIIS-----QDGSIRLPGVLFQKCDRE-----	-----EEIVVSKAGRLH--Q-----	-----EEIVVSKAGRLH--Q-----	-----TGKNAMLIDS
CnRrp4	-----MPTLILPGETVPIPS-----TSKAVVLPGISQSTPTLQASSSTAGSSSSAQIATRLGLH--SGKG--KERSQKLVIEG-----	-----TSKAVVLPGISQSTPTLQASSSTAGSSSSAQIATRLGLH--SGKG--KERSQKLVIEG-----	-----TSKAVVLPGISQSTPTLQASSSTAGSSSSAQIATRLGLH--SGKG--KERSQKLVIEG-----	-----TSKAVVLPGISQSTPTLQASSSTAGSSSSAQIATRLGLH--SGKG--KERSQKLVIEG-----
AtRrp4	-----MAAKVSVSPTSLNLIQVVPDVLDSLMN-----NQNTIKLGGSLQDNDA-----	-----ISAMRAGKLSYSKPN-----	-----ISAMRAGKLSYSKPN-----	-----KYWLES
OsRrp4	-----MESRKPFPFSAIVNDVNIHVPDVLDAEM-----TNQTIKLAGLRQDCDT-----	-----IQATSAGRRLSKPFS-----	-----IQATSAGRRLSKPFS-----	-----KYWLES
ScRrp4	-----NMMSQSQKIVLQPSRIVPGVLLAEGE-----FQIPWSPYILKINSK-----	-----YYSVTGKLVFDKDTQ-----	-----YYSVTGKLVFDKDTQ-----	-----FEVIFL
consensus	-----	g g	g g	g g
HeXOSC2	KT-RYIEGVGDIVVGRITVEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--GQGVLVQ--VSPSLVK-----	-----RQKTHFHDLPL-----	-----CGASVIL--NNGFIWII
MmRrp4	KT-RYIEGVGDIVVGRITVEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--GQGVLVQ--VSPSLVK-----	-----RQKTHFHDLPL-----	-----CGASVIL--NNGFIWII
DrRrp4	KT-RYIEGVGDIVVGRITVEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--GQGVLVQ--VSPSLVK-----	-----RQKTHFHDLPL-----	-----CGASVIL--NNGFIWII
DmRrp4	KS-RYVAGEVGDVVVARVSEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----YTRSLKYGKL--SQGLLVK--VFPALVK-----	-----RKKHFFHTPL-----	-----CGASVIL--NNGFIWII
CeRrp4	KQ-RYVAGEVGDVVVARVSEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRNNKYGKL--QQGLILK--VFPALVK-----	-----RKKHFFHTPL-----	-----YGMAVI--IGCNGVWVT
ScRrp4	KG-RYVAGEVGDVVVARVSEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--RNGMFCQ--VPSLIV-----	-----RAKHTHTPL-----	-----GNITVVLGVNGVWVT
SpRrp4	RS-KYVPEIGDILIGIKIAEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--RNGVFLK--VPPALV-----	-----RSKSHAYALA-----	-----GGVDI--ILSNGVWVT
CnRrp4	RS-KYVPEIGDILIGIKIAEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----HTRSLKYGKL--RNGVFLK--VPPALV-----	-----RSKSHAYALA-----	-----GGVDI--ILSNGVWVT
AtRrp4	RA-RYKPEVGDIVVGRITVEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----QARSQRYGKL--EKGQLLK--VDPYLVK-----	-----RSKSHAYALA-----	-----SLGDI--LIGCNGFVWVG
OsRrp4	RA-RYKPEVGDIVVGRITVEVQQRKRWKVTNSR-LDSVLLSSNNLPGGELRRRSAAEDELAMRGFLOEGLLISAEVQAVFDGAVSL-----	-----QARSQRYGKL--EKGQLLK--VDPYLVK-----	-----RSKSHAYALA-----	-----SLGDI--LIGCNGFVWVG
HeXOSC3	QQRKRVVPGKDHVIGIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
MmRrp4	QQRKRVVPGKDHVIGIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
DrRrp4	QQRKRVVPGKDHVIGIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
DmRrp4	QQRKRVVPGKDHVIGIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
CeRrp4	HSKRYIPQEGDRVIAIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
ScRrp4	SSKRYIPQEGDRVIAIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
SpRrp4	TRKRYIPATNDQVIGIISRAEGYRVLDIGSS-QMA--QLDALAPEGATKRS-----RPNLQVGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
CnRrp4	NSKRYIPQEGDRVIAIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
AtRrp4	SHKRYIPQEGDRVIAIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
OsRrp4	SKKRYIPQEGDRVIAIVAKSGDIFKVDVGGG-EPA--SLSYLAPEAGTKRN-----RPNVQGDVLYGQCVANAKMPEMVCIDS--CGRANGMVIQDCLLFK-VTLGLIRKLLA--	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK	-----PDCEII--QEVGLKVLPLEIVFGMNGR--VWK
ScRrp4	EGSFYYPKINDIVIGLVDEVIYGVVVDIKGP-YKAYLPAS-----NLLGRSINVGDELARLVLDVYGIARIDENFDSIDFVL-----SVRGKDLGRV--SNGVYFD--	-----IMPVKVPRVIGNKRSMYTLI-----	-----SKGSGSIFVANNR--WAT	-----SKGSGSIFVANNR--WAT
consensus	rY d v i g v f d 1 1 a r	g g i g m r r g i g	i	i v g N G I W
HeXOSC2	PTPEHKE-----EE-AGGFIANL-EPV-SLADREIVSRRLNRCIISLVTQRMLYDTSILYCYEASLPHQIKDI-----	-----LKPEIME-EIVMETRQRL-----	-----LEQEG--	-----LEQEG--
MmRrp4	PTPEHKE-----EE-AGGFIANL-EPV-SLADREIVSRRLNRCIISLVTQRMLYDTSILYCYEASLPHQIKDI-----	-----LKPEIME-EIVMETRQRL-----	-----LEQEG--	-----LEQEG--
DrRrp4	PTPALQE-----EQ-AGGFYTM-EPV-SLSDREIVSRRLNRCIISLVTQRMLYDTSILYCYEASLPHQIKDI-----	-----LKPEIME-EIVLETRQRL-----	-----MEHEG--	-----MEHEG--
DmRrp4	PTPQOEE-----EGEGGFAQNLNHEV-PRADREIVSRRLNRCIISLVTQRMLYDTSILYCYEASLPHQIKDI-----	-----LQONAIY-DIQOQOARL-----	-----RDADE--	-----RDADE--
CeRrp4	PSLPETTL-----EGRS--HVH-EQIVPPDVTMVRVIAACVRLLRDYSIYLNLSLTCYEMSQPYEIKELAEQDTSRRLAYL--	-----IAARLQLQEQ-----	-----QKQ--	-----QKQ--
ScRrp4	KTSQMDLARDTPSANNSSSIKSTGPTGAVSLNPSITRLEEESSQIYSDENDPSISNNIRQAICRYANVIRALAFCEIGITQQRIVSAYEASMYV--	-----SNVDEL--	-----IEKNVM-ESIGSDILTAEK-----	-----MRGNNG--
SpRrp4	KHNENQI-----SSVITRLEEESEIYSNENDEI-DGTRLANISRVSTCKLASRSLDPTQASITNFEESLVP--	-----SNLQDL--	-----TVPRM--DQIAMEAMQ-----	-----TVPRM--DQIAMEAMQ-----
CnRrp4	NGTSQARE-----GGGDFEGYNSQADEI-PQGRBAILVAMIKKALASGVPTETLIGESVYAKENVPSGSGPFDSEBESVMEVGLFLEFA-----	-----	-----	-----
AtRrp4	EHEVYRDPMAIDD-----QKDEMISSS-STGKQSHIPLERQICRIGANRVLNLSGTVTVLEVMETVMSNSKNIDH-----	-----DMLGSEFHVVAAEAEARRRTRK--	-----	-----
OsRrp4	EHHVVGENANMEE-----NKL-----N-LSAEVENFTPLERTRKHCRLANAVRSLALGTLTVLEIIETAEASVSNIEIN-----	-----NMLGAEFVQTAHEVYKRRADLRLKSSGAR-----	-----	-----
HeXOSC3	AKTIQOTLILANILEACEHMTSQRKQ-----IFSRLES-----	-----IFARLAE-----	-----	-----
MmRrp4	AKTIQOTLILANILEACEHMTSQRKQ-----IFSRLES-----	-----IFARLAE-----	-----	-----
DrRrp4	AKTIQOTLILANILEACEHMTSQRKQ-----IFSRLES-----	-----IFARLAE-----	-----	-----
DmRrp4	AHSLKETVALANASALEQSGCAEIDK-----LVQNNYTRVSD-----	-----IAKRQFEKILTVKEE-----	-----	-----
CeRrp4	ASTSDDIKIHDIINKSEFITDEDELIT-----LVQNNYTRVSD-----	-----IAKRQFEKILTVKEE-----	-----	-----
ScRrp4	CELSNLTACYRTIMECCQKNDTAAFDK-----IAKRQFEKILTVKEE-----	-----IAKRQFEKILTVKEE-----	-----	-----
SpRrp4	SENLSLTVLICTAIRNCFMSEDEEQIKY-----CKDLIKL-----	-----CKDLIKL-----	-----	-----
CnRrp4	AETVGETLAMKRLIESVDHIIIVVEKA-----IKEKREYS-----	-----IKEKREYS-----	-----	-----
AtRrp4	ATAPRIVLIVANAIMNSESLSGTQRI-----MVEKLLAISD-----	-----MVEKLLAISD-----	-----	-----
OsRrp4	AFSPNVIIVSNALIKESLSGIGQRS-----RKI-----	-----ENESH-----	-----IKGLDRIKQFIEKGLGRNASSGETKTNS-----	-----IKGLDRIKQFIEKGLGRNASSGETKTNS-----
ScRrp4	CFSRPFSEELIEAI-----RKI-----	-----ENESH-----	-----IKGLDRIKQFIEKGLGRNASSGETKTNS-----	-----IKGLDRIKQFIEKGLGRNASSGETKTNS-----
consensus				

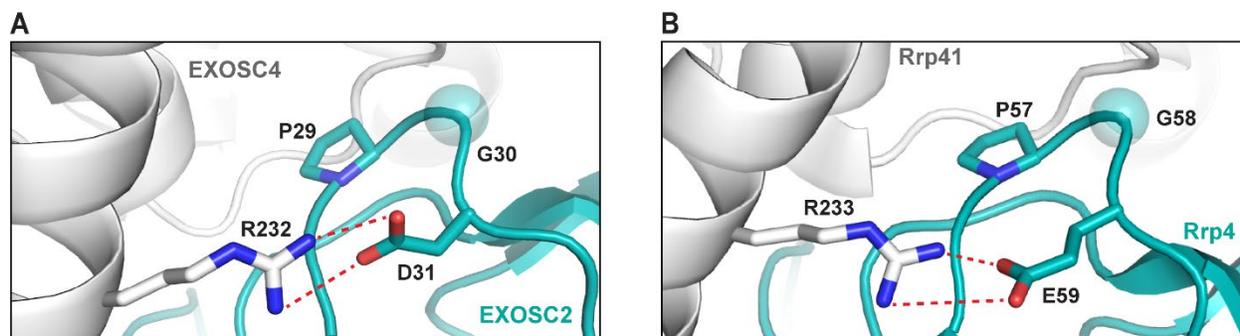
Supplementary Figure S1. Protein sequence alignment of human EXOSC2 and EXOSC3.

Protein sequence alignment of the human RNA exosome cap subunits EXOSC2 and EXOSC3, and *S.*

cerevisiae cap subunits Rrp4 and Rrp40, including other EXOSC2/Rrp4 and EXOSC3/Rrp40 orthologs.

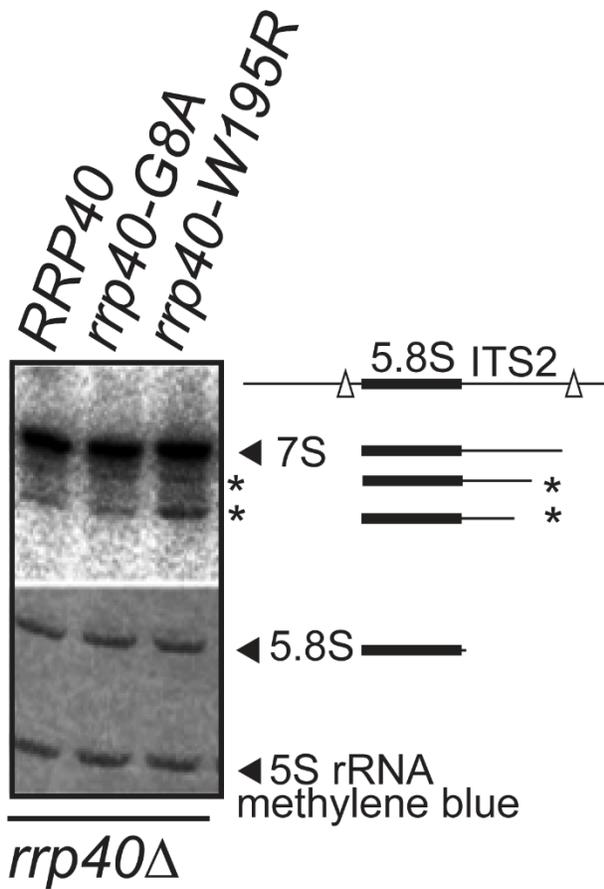
Conserved EXOSC2 amino acids substituted in patients with SHRF (Short stature, Hearing loss, Retinitis pigmentosa and distinctive Facies) syndrome [40] and EXOSC3 amino acids substituted in patients with

PCH1b (pontocerebellar hypoplasia type 1b) disease [43] are highlighted in green. Identical residues (in red) and similar residues (in blue) are indicated. The different species aligned are as follows: Hs: *Homo sapiens*, Mm: *Mus musculus*, Dr: *Danio rerio*, Dm: *Drosophila melanogaster*, Ce: *Caenorhabditis elegans*, Sc: *Saccharomyces cerevisiae*, Sp: *Schizosaccharomyces pombe*, Cn: *Cryptococcus neoformans*, At: *Arabidopsis thaliana*, Os: *Oryza sativa*, Ss: *Sulfolobus solfataricus* (archaea).



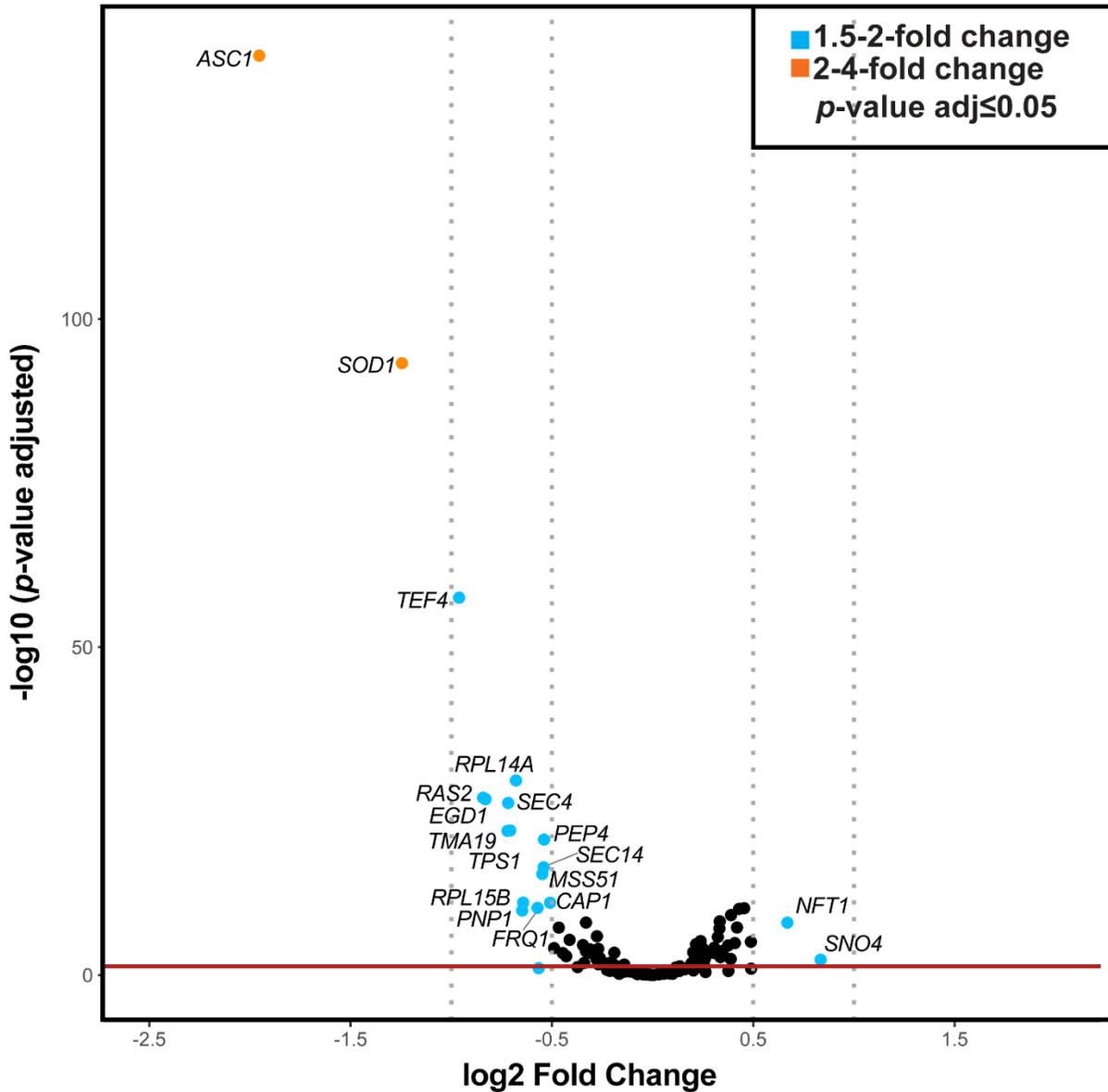
Supplementary Figure S2. Modeling of the human EXOSC2-EXOSC4 and yeast Rrp4-Rrp41 interface show structural conservation.

(A) Zoomed-in representations of the interface between the human RNA exosome cap subunit EXOSC2 (teal blue) and core subunit EXOSC4 (light gray). EXOSC2 residue Gly30 (G30) facilitates a β turn that positions EXOSC2 residue Asp31 (D31) near an arginine in EXOSC4, Arg232 (R232). Structural modeling shows a salt bridge that forms between EXOSC2 D31 and EXOSC4 R232, represented by the red dashed lines. (B) Zoomed-in representation of the interface between the yeast RNA exosome cap subunit Rrp4 (teal blue) and core subunit Rrp41 (light gray). Rrp4 residue Gly58 (G58), which corresponds to EXOSC2 G30, facilitates a β turn that positions Rrp4 Glu59 (E59) near an arginine in the EXOSC4 yeast ortholog, Rrp41, Arg233 (R233). Structural modeling shows a salt bridge forms between Rrp4 E59 and Rrp41 R233, represented by the red dashed lines. Structural modeling in (A) was performed with the human RNA exosome structure (PDB 6D6R) [28] and in (B) with the yeast RNA exosome structure (PDB 6FSZ) [29] using PyMOL [75].



Supplementary Figure S3. Increased input signal levels for rRNA northern blot in Figure 5A emphasize previously observed accumulation of 5.8S precursors in *rrp40-W195R* cells.

The lanes for the *RRP40*, *rrp40-G8A* and *rrp40-W195R* samples from the northern blot displayed in Figure 5A are shown here with the input signal levels increased. Total RNA from *RRP40*, *rrp40-G8A*, and *rrp40-W195R* cells grown at 37°C was analyzed by northern blotting with a 5.8S-ITS2 probe to detect 7S pre-rRNA. Mature 5.8S rRNA and 5S rRNA was detected by methylene blue staining as a loading control. The simplified schematics to the right illustrate the processing steps of 7S rRNA precursor following endonucleolytic cleavage from larger 27S precursor (indicated by white triangles). Accumulation of 5.8S precursors is evident in *rrp40-W195R* (labeled with asterisks).



Supplementary Figure S4. Volcano plot of autophagy transcripts differentially expressed in the *rrp4-G226D* RNA-Seq.

A total of 18 autophagy transcripts show ± 1.5 -fold change (p -value adjusted < 0.05) in *rrp4-G226D* cells compared to *RRP4* cells. Of those, 2 transcripts are increased 1.5-2 fold (*NFT1* and *SNO4* (blue)) and 16 transcripts are decreased 1.5-2-fold (*TEF4*, *RPL14A*, *RAS2*, *EGD1*, *SEC4*, *TPS1*, *TMA19*, *PEP4*, *SEC14*, *MSS51*, *RPL15B*, *CAP1*, *FRQ1*, *PNP1* (blue)) with two transcripts decreased > 2 -fold (*ASC1* and *SOD1* (orange)).

Table S1. *S. cerevisiae* Strains and Plasmids used in this study.

Yeast Strains and Plasmids

Strain/Plasmid	Description	Source
<i>rrp4Δ</i> (yAV1103)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, lys2Δ0, rrp4Δ::NEO, [RRP4, URA3]</i>	(LOSH 2018)
<i>rrp40Δ</i> (yAV1107)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, rrp40Δ::NEO, [RRP40, URA3]</i>	(SCHAEFFER <i>et al.</i> 2009)
<i>rrp4Δ mpp6Δ</i> (ACY2471)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, lys2Δ0, rrp4Δ::NEO, mpp6Δ::natMX4, [RRP4, URA3]</i>	This study
<i>rrp4Δ rrp47Δ</i> (ACY2474)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, lys2Δ0, rrp4Δ::NEO, rrp47Δ::natMX4, [RRP4, URA3]</i>	This study
<i>rrp4Δ rrp6Δ</i> (ACY2478)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, lys2Δ0, rrp4Δ::NEO, rrp6Δ::natMX4, [RRP4, URA3]</i>	This study
<i>rrp40Δ mpp6Δ</i> (ACY2638)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, rrp40Δ::NEO, mpp6Δ::natMX4, [RRP40, URA3]</i>	This study
<i>rrp40Δ rrp47Δ</i> (ACY2462)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, rrp40Δ::NEO, rrp47Δ::natMX4, [RRP40, URA3]</i>	This study
<i>rrp40Δ rrp6Δ</i> (ACY2466)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, rrp40Δ::NEO, rrp6Δ::natMX4, [RRP40, URA3]</i>	This study
<i>RRP43-TAP</i> (ACY2788)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, met15Δ0, RRP43-TAP:HIS3MX6</i>	(GHAEMMAGHAM <i>et al.</i> 2009)
<i>RRP43-TAP rrp4Δ</i> (ACY2803)	<i>MATa, ura3Δ0, leu2Δ0, his3Δ1, met15Δ0, RRP43-TAP:HIS3MX6, rrp4Δ::neoMX, [RRP4, URA3]</i>	This study
pRS315	<i>CEN6, LEU2, amp^R</i>	(SIKORSKI AND HIETER 1989)
pRS313	<i>CEN6, HIS3, amp^R</i>	(SIKORSKI AND HIETER 1989)
pAC3161	<i>RRP40-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	(FASKEN <i>et al.</i> 2017)
pAC3162	<i>rrp40-G8A-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	(FASKEN <i>et al.</i> 2017)
pAC3259	<i>rrp40-W195R-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	(FASKEN <i>et al.</i> 2017)
pAC3652	<i>RRP40-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3655	<i>rrp40-W195R-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3474	<i>RRP4-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3476	<i>rrp4-G58V-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3477	<i>rrp4-G226D-2xMyc in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3656	<i>RRP4-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3659	<i>rrp4-G226D-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3669	<i>RRP4-2xMyc-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3670	<i>rrp4-G58V-2xMyc-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC3672	<i>rrp4-G226D-2xMyc-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This study
pAC2897	<i>MTR4, 2μ, URA3, amp^R</i>	(FASKEN <i>et al.</i> 2011)
pAC3713	<i>MTR4, RRP40, CEN6, URA3, amp^R</i>	This study
pAC3714	<i>MTR4, RRP4, CEN6, URA3, amp^R</i>	This study
pAC3719	<i>MTR4-2xFLAG-Native 3'UTR in pRS313, CEN6, HIS3, amp^R</i>	This study
pAC4096	<i>MTR4-Native 3'UTR in pRS313, CEN6, HIS3, amp^R</i>	This study
pAC4099	<i>mtr4-F7A-F10A-Native 3'UTR in pRS313, CEN6, HIS3, amp^R</i>	This study

Table S2. DNA Oligonucleotides employed for RT-qPCR

DNA Oligonucleotides used for Quantitative RT-PCR

Description	Sequence (5'-3')	Name
pre- <i>U4</i> snRNA Fwd	ATCCTTATGCACGGGAAATACG	AC5722
pre- <i>U4</i> snRNA Rev	AAAGAATGAATATCGGTAATG	AC5723
<i>U14</i> snoRNA (snR128) Fwd	GATCACGGTGATGAAAGACTGG	AC5397
<i>U14</i> snoRNA (snR128) Rev	CTACAGTATACGATCACTCAGACATCCTA	AC5398
<i>snR44</i> snoRNA Fwd	GCATTTCCACATGGGATTTAA	AC6270
<i>snR44</i> snoRNA Rev	ATGGTGTGATCGGGCAGTAT	AC6272
<i>TLC1</i> ncRNA Fwd	AAGGCAAGGGTGTCTTTCT	AC6420
<i>TLC1</i> ncRNA Rev	TTCCGCTTGGAAAATAATGC	AC6421
pre- <i>TLC1</i> ncRNA Fwd	GTATTGTAGAAATCGCGCGTAC	AC7593
pre- <i>TLC1</i> ncRNA Rev	CCGCCTATCCTCGTCATGAAC	AC7594
<i>RPS3</i> mRNA Fwd	TCCAACCAAGACCGAAGTTATC	AC9226
<i>RPS3</i> mRNA Rev	GTACCTGGAGCGTACTTGAATC	AC9227
<i>RPL15A</i> mRNA Fwd	CCAGACAAGGCTAGAAGATTGG	AC9309
<i>RPL15A</i> mRNA Rev	CCGTAAGTAGCACCCCTTTGG	AC9308
<i>INO1</i> mRNA Fwd	TTGGACTGCAAATACTGAGAGG	AC9303
<i>INO1</i> mRNA Rev	AAGATCGTGGAAAGGAGCAATC	AC9302
<i>PTH4</i> mRNA Fwd	ACTGTGCTTGGATTCCCTCAG	AC9248
<i>PTH4</i> mRNA Rev	CTATAGAATCGCTGCCCTTAGC	AC9249
<i>HXK2</i> mRNA Fwd	TACTGGTGTCAATGGTGCTTAC	AC9307
<i>HXK2</i> mRNA Rev	TTGGAGCAGATGGTGGAAATG	AC9306
<i>TDH1</i> mRNA Fwd	GGTAGATACAAGGGTACTGTTTCC	AC9230
<i>TDH1</i> mRNA Rev	TGAGCGGTGTCCAATTCC	AC9232
<i>CUT501</i> ncRNA Fwd	GGTTCAACGTTGCAGGATCT	AC9254
<i>CUT501</i> ncRNA Rev	GCTAGCACCTGTTGCTGTAAT	AC9255
<i>CUT770</i> ncRNA Fwd	AAACAACCCGCTAGTGTGAC	AC9262
<i>CUT770</i> ncRNA Rev	AGAGCAACTCACTGCAAAGG	AC9263
<i>CUT896</i> ncRNA Fwd	ATCAGCAGGTGTCATGTTACAG	AC9256
<i>CUT896</i> ncRNA Rev	CCCAGAGGCAAAGATGTTAAGT	AC9257
<i>NRD1</i> mRNA Fwd	CAAGCAGAGGTGCAAACAAATC	AC9244
<i>NRD1</i> mRNA Rev	GCTGGATCTGTGGAAGTCAA	AC9245
<i>NAB3</i> mRNA Fwd	ACAGTTCGGTAGGCTCAGATAG	AC9246
<i>NAB3</i> mRNA Rev	GGCGAAGTTCGACCTCTTTATC	AC9247
<i>ALG9</i> mRNA Fwd	CACGGATAGTGGCTTTGGTGAACAATTAC	AC5067
<i>ALG9</i> mRNA Rev	TATGATTATCTGGCAGCAGGAAAGAACTTGGG	AC5068

Table S3. Summary of in silico predictions for pathogenic amino acid substitutions in EXOSC2 and Rrp4.

Exosomopathy amino acid substitution	<i>Hs</i> EXOSC2		<i>Sc</i> Rrp4	
	G30V	G198D	G58V	G226D
Missense 3D* (Phyre2)	Neutral	Damaging	Neutral	Damaging
PolyPhen-2** (HumDiv)	Possibly Damaging (Score: 1.000)	Probably Damaging (Score: 1.000)	Probably Damaging (Score: 1.000)	Probably Damaging (Score: 1.000)
Provean†	Deleterious (-7.938)	Deleterious (-6.35)	Deleterious (-8.981)	Deleterious (-6.517)
SNAP2‡	Effect (Score: 91)	Effect (Score: 94)	Effect (Score: 63)	Effect (Score: 92)

***Missense 3D** part of tool suite **Phyre2** (**Protein Homology/analogY Recognition Engine**) v2.0; predicts structural changes introduced by an amino acid substitution through three-dimensional protein modeling.

****PolyPhen-2** (**Polymorphism Phenotyping**) v2 HumDiv trained model; Naive Bayes posterior probability that mutation is damaging and qualitative classification as benign, possibly damaging, or probably damaging based on 5%/10% false positive rate (FPR) thresholds (FPR, the chance the mutation is classified as damaging when it is non-damaging).

†**PROVEAN** (**Protein Variation Effect Analyzer**) v1.1; delta alignment scores equal to or below predefined threshold (-2.5), protein variant predicted to have "deleterious" effect on function.

‡**SNAP2** predicted score for functional effects of mutations; scores range from -100 strong neutral prediction to +100 strong effect prediction.

2.9 References

1. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'->5' exoribonucleases*. Cell, 1997. **91**(4): p. 457-466.
2. Mitchell, P., E. Petfalski, and D. Tollervy, *The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism*. Genes Dev, 1996. **10**(4): p. 502-13.
3. Hou, D., M. Ruiz, and E.D. Andrulis, *The ribonuclease Dis3 is an essential regulator of the developmental transcriptome*. BMC Genomics, 2012. **13**.
4. Lim, S.J., et al., *Genome-wide localization of exosome components to active promoters and chromatin insulators in Drosophila*. Nucleic Acids Res, 2013. **41**(5): p. 2963-80.
5. Pefanis, E., et al., *Noncoding RNA transcription targets AID to divergently transcribed loci in B cells*. Nature, 2014. **514**(7522): p. 389-93.
6. Lorentzen, E., et al., *RNA channelling by the archaeal exosome*. Embo Reports, 2007. **8**(5): p. 470-476.
7. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 227-239.
8. Fasken, M.B., et al., *The RNA Exosome and Human Disease*. Methods Mol Biol, 2020. **2062**: p. 3-33.
9. Allmang, C., et al., *Functions of the exosome in rRNA, snoRNA and snRNA synthesis*. Embo Journal, 1999. **18**(19): p. 5399-5410.
10. van Hoof, A., P. Lennertz, and R. Parker, *Yeast Exosome Mutants Accumulate 3'-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs*. Molecular and Cellular Biology, 2000. **20**(2): p. 441-452.
11. Parker, R., *RNA Degradation in Saccharomyces cerevisiae*. Genetics, 2012. **191**(3): p. 671-702.
12. Schneider, C., et al., *Transcriptome-wide Analysis of Exosome Targets*. Molecular Cell, 2012. **48**(3): p. 422-433.
13. Wyers, F., et al., *Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase*. Cell, 2005. **121**(5): p. 725-737.
14. Moore, M.J. and N.J. Proudfoot, *Pre-mRNA processing reaches back to transcription and ahead to translation*. Cell, 2009. **136**(4): p. 688-700.
15. Schneider, C. and D. Tollervy, *Threading the barrel of the RNA exosome*. Trends in Biochemical Sciences, 2013. **38**(10): p. 485-493.
16. Makino, D.L., M. Baumgaertner, and E. Conti, *Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex*. Nature, 2013. **495**(7439): p. 70-75.
17. Briggs, M.W., K.T. Burkard, and J.S. Butler, *Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation*. J Biol Chem, 1998. **273**(21): p. 13255-63.

18. Wasmuth, E.V., K. Januszyk, and C.D. Lima, *Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA*. *Nature*, 2014. **511**(7510): p. 435-9.
19. Bonneau, F., et al., *The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation*. *Cell*, 2009. **139**(3): p. 547-59.
20. Liu, Q., J.C. Greimann, and C.D. Lima, *Reconstitution, activities, and structure of the eukaryotic RNA exosome*. *Cell*, 2006. **127**(6): p. 1223-1237.
21. Zinder, J.C., E.V. Wasmuth, and C.D. Lima, *Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3*. *Mol Cell*, 2016. **64**(4): p. 734-745.
22. Zinder, J.C. and C.D. Lima, *Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors*. *Genes & Development*, 2017. **31**(2): p. 88-100.
23. de la Cruz, J., et al., *Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae*. *Embo j*, 1998. **17**(4): p. 1128-40.
24. LaCava, J., et al., *RNA degradation by the exosome is promoted by a nuclear polyadenylation complex*. *Cell*, 2005. **121**(5): p. 713-724.
25. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. *Cell Reports*, 2017. **20**(10): p. 2279-2286.
26. Schuch, B., et al., *The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase*. *Embo Journal*, 2014. **33**(23): p. 2829-2846.
27. Wasmuth, E.V., et al., *Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase*. *Elife*, 2017. **6**: p. 24.
28. Weick, E.M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. *Cell*, 2018. **173**(7): p. 1663-1677 e21.
29. Schuller, J.M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. *Science*, 2018. **360**(6385): p. 219-222.
30. Stuparevic, I., et al., *Cotranscriptional Recruitment of RNA Exosome Cofactors Rrp47p and Mpp6p and Two Distinct Trf-Air-Mtr4 Polyadenylation (TRAMP) Complexes Assists the Exonuclease Rrp6p in the Targeting and Degradation of an Aberrant Messenger Ribonucleoprotein Particle (mRNP) in Yeast*. *Journal of Biological Chemistry*, 2013. **288**(44): p. 31816-31829.
31. Rodríguez-Galán, O., et al., *Immature large ribosomal subunits containing the 7S pre-rRNA can engage in translation in Saccharomyces cerevisiae*. *RNA biology*, 2015. **12**(8): p. 838-846.
32. Vaňáčová, Š., et al., *A New Yeast Poly(A) Polymerase Complex Involved in RNA Quality Control*. *PLOS Biology*, 2005. **3**(6): p. e189.
33. Houseley, J. and D. Tollervey, *The nuclear RNA surveillance machinery: the link between ncRNAs and genome structure in budding yeast?* *Biochim Biophys Acta*, 2008. **1779**(4): p. 239-46.
34. Houseley, J. and D. Tollervey, *The many pathways of RNA degradation*. *Cell*, 2009. **136**(4): p. 763-76.

35. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nat Rev Mol Cell Biol, 2016. **17**(4): p. 227-39.
36. Lubas, M., et al., *Interaction profiling identifies the human nuclear exosome targeting complex*. Mol Cell, 2011. **43**(4): p. 624-37.
37. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. J Neurol, 2013. **260**(7): p. 1866-70.
38. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. Nature Communications, 2014. **5**.
39. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
40. Di Donato, N., et al., *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*. Journal of Medical Genetics, 2016. **53**(6): p. 419-425.
41. Eggens, V.R.C., et al., *EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations*. Orphanet Journal of Rare Diseases, 2014. **9**(1): p. 23.
42. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
43. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nature Genetics, 2012. **44**(6): p. 704-U134.
44. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. Hum Mol Genet, 2020. **29**(13): p. 2218-2239.
45. Morton, D.J., et al., *The RNA exosome and RNA exosome-linked disease*. Rna, 2018. **24**(2): p. 127-142.
46. Burns, D., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. Neuromuscul Disord, 2017. **27**: p. S38.
47. Yang, X., et al., *Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis*. Human Molecular Genetics, 2019. **29**(4): p. 541-553.
48. Oddone, A., et al., *Structural and biochemical characterization of the yeast exosome component Rrp40*. EMBO reports, 2007. **8**(1): p. 63-69.
49. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. Genetics, 2017. **205**(1): p. 221-+.
50. Morton, D.J., et al., *A Drosophila model of Pontocerebellar Hypoplasia reveals a critical role for the RNA exosome in neurons*. PLoS Genet, 2020. **16**(7): p. e1008901.

51. Gillespie, A., et al., *Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in S. cerevisiae*. RNA, 2017. **23**(4): p. 466-472.
52. de Amorim, J., Slavotinek, A., Fasken, M.B., Corbett, A.H., Morton, D.J., *Modeling Pathogenic Variants in the RNA Exosome*. RNA & Disease, 2020. **7**.
53. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae*. Genetics, 1989. **122**(1): p. 19-27.
54. Halevy, A., et al., *Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia*. J Neurol, 2014. **261**(11): p. 2165-9.
55. Coy, S., et al., *The Sm complex is required for the processing of non-coding RNAs by the exosome*. PloS one, 2013. **8**(6): p. e65606-e65606.
56. Donahue, T.F. and S.A. Henry, *myo-Inositol-1-phosphate synthase. Characteristics of the enzyme and identification of its structural gene in yeast*. J Biol Chem, 1981. **256**(13): p. 7077-85.
57. Klig, L.S. and S.A. Henry, *Isolation of the yeast INO1 gene: located on an autonomously replicating plasmid, the gene is fully regulated*. Proc Natl Acad Sci U S A, 1984. **81**(12): p. 3816-20.
58. Delan-Forino, C., C. Schneider, and D. Tollervy, *Transcriptome-wide analysis of alternative routes for RNA substrates into the exosome complex*. PLOS Genetics, 2017. **13**(3): p. e1006699.
59. Kuleshov, M.V., et al., *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update*. Nucleic Acids Res, 2016. **44**(W1): p. W90-7.
60. Kuleshov, M.V., et al., *modEnrichr: a suite of gene set enrichment analysis tools for model organisms*. Nucleic Acids Research, 2019. **47**(W1): p. W183-W190.
61. Chen, E.Y., et al., *Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool*. BMC Bioinformatics, 2013. **14**: p. 128.
62. Belair, C., S. Sim, and S.L. Wolin, *Noncoding RNA Surveillance: The Ends Justify the Means*. Chem Rev, 2018. **118**(8): p. 4422-4447.
63. Steinmetz, E.J. and D.A. Brow, *Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association*. Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6699-704.
64. Steinmetz, E.J., et al., *RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts*. Nature, 2001. **413**(6853): p. 327-31.
65. Wolin, S.L., S. Sim, and X. Chen, *Nuclear noncoding RNA surveillance: is the end in sight?* Trends Genet, 2012. **28**(7): p. 306-13.
66. Mitchell, P., et al., *Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs*. Mol Cell Biol, 2003. **23**(19): p. 6982-92.
67. Davis, C.A. and M. Ares, Jr., *Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3262-7.

68. Marquardt, S., D.Z. Hazelbaker, and S. Buratowski, *Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species*. *Transcription*, 2011. **2**(3): p. 145-154.
69. Xu, Z., et al., *Bidirectional promoters generate pervasive transcription in yeast*. *Nature*, 2009. **457**(7232): p. 1033-7.
70. Yatsuka, H., et al., *Exosc2 deficiency leads to developmental disorders by causing a nucleotide pool imbalance in zebrafish*. *Biochemical and Biophysical Research Communications*, 2020.
71. Belair, C., et al., *The RNA exosome nuclease complex regulates human embryonic stem cell differentiation*. *J Cell Biol*, 2019.
72. Gudipati, R.K., et al., *Extensive Degradation of RNA Precursors by the Exosome in Wild-Type Cells*. *Molecular Cell*, 2012. **48**(3): p. 409-421.
73. Fox, M.J., et al., *The Exosome Component Rrp6 Is Required for RNA Polymerase II Termination at Specific Targets of the Nrd1-Nab3 Pathway*. *PLOS Genetics*, 2015. **11**(2): p. e1004999.
74. Adams, A., et al., *Methods in Yeast Genetics*. 1997, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
75. PyMOL, *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC*.
76. Rodrigues, C.H.M., et al., *mCSM-PPI2: predicting the effects of mutations on protein-protein interactions*. *Nucleic Acids Research*, 2019. **47**(W1): p. W338-W344.
77. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations*. *Nat Methods*, 2010. **7**(4): p. 248-9.
78. Choi, Y., et al., *Predicting the Functional Effect of Amino Acid Substitutions and Indels*. *PLOS ONE*, 2012. **7**(10): p. e46688.
79. Choi, Y., *A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein*, in *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine*. 2012, Association for Computing Machinery: Orlando, Florida. p. 414-417.
80. Hecht, M., Y. Bromberg, and B. Rost, *Better prediction of functional effects for sequence variants*. *BMC Genomics*, 2015. **16 Suppl 8**(Suppl 8): p. S1.
81. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*. Second ed. 1989, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
82. Schaeffer, D., et al., *The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities*. *Nat Struct Mol Biol*, 2009. **16**(1): p. 56-62.
83. Losh, J., *Identifying Subunit Organization and Function of the Nuclear RNA Exosome Machinery*. 2018, University of Texas: UT GSBS Dissertations and Theses.

84. Ghaemmaghami, S., et al., *Global analysis of protein expression in yeast*. Nature, 2003. **425**(6959): p. 737-41.
85. Fasken, M.B., et al., *Air1 zinc knuckles 4 and 5 and a conserved IWRXY motif are critical for the function and integrity of the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) RNA quality control complex*. J Biol Chem, 2011. **286**(43): p. 37429-45.
86. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
87. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics, 2012. **29**(1): p. 15-21.
88. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features*. Bioinformatics, 2014. **30**(7): p. 923-30.
89. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biology, 2014. **15**(12): p. 550.
90. Cherry, J.M., et al., *Saccharomyces Genome Database: the genomics resource of budding yeast*. Nucleic Acids Res, 2012. **40**(Database issue): p. D700-5.

Chapter III: *In vivo* Characterization of the Critical Interaction between the RNA Exosome and the Essential RNA Helicase Mtr4 in *Saccharomyces cerevisiae*

The research and data presented in this chapter is published in *G3* as “*In vivo* Characterization of the Critical Interaction between the RNA Exosome and the Essential RNA Helicase Mtr4 in *Saccharomyces cerevisiae*”

Sterrett, M. C., D. Farchi, S. E. Strassler, L. H. Boise, M. B. Fasken and A. H. Corbett (2022). "*In vivo* Characterization of the Critical Interaction between the RNA Exosome and the Essential RNA Helicase Mtr4 in *Saccharomyces cerevisiae*." *G3 Genes|Genomes|Genetics*, 2023;, jkad049, <https://doi.org/10.1093/g3journal/jkad049>.

Author Contributions:

M.C. Sterrett created all the data for published figures, compiled, and wrote the manuscript. The original basis for this work comes from D. Farchi’s honors thesis which was designed by **M.C. Sterrett**. **M.C. Sterrett** also served as D. Farchi’s lab mentor. S. E. Strassler assisted in modeling the amino acid substitutions. L.H. Boise originally identified the human *EXOSC2* mutation and aided in analysis of the CoMMpass data.

3.1 Abstract

The RNA exosome is a conserved molecular machine that processes/degrades numerous coding and non-coding RNAs. The 10-subunit complex is composed of three S1/KH cap subunits (human EXOSC2/3/1; yeast Rrp4/40/Csl4), a lower ring of six PH-like subunits (human EXOSC4/7/8/9/5/6; (yeast Rrp41/42/43/45/46/Mtr3), and a singular 3'-5' exo/endonuclease DIS3/Rrp44. Recently, several disease-linked missense mutations have been identified in structural cap and core RNA exosome genes. In this study, we characterize a rare multiple myeloma patient missense mutation that was identified in the cap subunit gene *EXOSC2*. This missense mutation results in a single amino acid substitution, p.Met40Thr, in a highly conserved domain of EXOSC2. Structural studies suggest this Met40 residue makes direct contact with the essential RNA helicase, MTR4, and may help stabilize the critical interaction between the RNA exosome complex and this cofactor. To assess this interaction *in vivo*, we utilized the *Saccharomyces cerevisiae* system and modeled the *EXOSC2* patient mutation into the orthologous yeast gene *RRP4*, generating the variant *rrp4 M68T*. The *rrp4 M68T* cells show accumulation of certain RNA exosome target RNAs and show sensitivity to drugs that impact RNA processing. We also identified robust negative genetic interactions between the *rrp4 M68T* variant and *mtr4* mutant variants. A biochemical approach revealed that Rrp4 M68T has decreased increased with Mtr4, consistent with these genetic results. This study suggests that the *EXOC2* mutation identified in a multiple myeloma patient impacts the function of the RNA exosome and provides an *in vivo* assessment of a critical interface between the RNA exosome and Mtr4.

3.2 Introduction

The RNA exosome is a highly conserved exo/endonuclease complex that has an essential role in 3' to 5' processing and degradation of nearly every species of RNA [1, 2]. First identified in *Saccharomyces cerevisiae* in a screen for ribosomal RNA processing (*rrp*) mutants [3, 4], the RNA exosome is essential in all organisms studied thus far [4-8]. In addition to ribosomal RNA precursors, the RNA exosome processes a variety of small non-coding RNAs (ncRNAs), including small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) [9-12]. The RNA exosome also plays roles in targeting RNA for degradation and decay, including non-functional or aberrant mRNAs and nuclear transcripts that result from pervasive transcription such as cryptic unstable transcripts (CUTs) in budding yeast or promoter upstream transcripts (PROMPTs) in humans [13-17]. The RNA exosome complex is composed of a 9-subunit structural core and a single exo/endonuclease [DIS3/DIS3L (human); Dis3/Rrp44 (budding yeast)]. As shown in Figure 1A, the 9-subunit structural core is composed of three S1/KH cap subunits (EXOSC1/2/3; Csl4/Rrp4/Rrp40) and a lower ring of six PH-like subunits (EXOSC4/5/6/7/8/9; Rrp41/Rrp46/Mtr3/Rrp42/Rrp43/Rrp45). The nuclear RNA exosome has an additional 3'-5' exonuclease, EXOSC10/Rrp6, that associates with the complex and aids in nuclear RNA targeting and processing [18, 19]. Structural studies demonstrate that the overall organization of the RNA exosome is conserved (Figure 1B) suggesting not only evolutionary conservation of the RNA exosome function but structure as well [20-22]. The vast array of targets and evolutionary conservation of the complex components indicates a fundamental role of the RNA exosome in several cellular processes, including but not limited to, maintaining genome integrity, translation, and cell differentiation through degradative and surveillance pathways [23].

RNA exosome specificity for a broad set of target transcripts is conferred in part through interactions with cofactor proteins, which aid the RNA exosome in target recognition, RNA unwinding, degradation, and catalysis in both the nucleus and the cytoplasm [1, 12]. Many nuclear RNA exosome cofactors were first characterized in budding yeast, including the Rrp6 obligate binding partner Rrp47, Mpp6 and the essential 3' to 5' DExH box RNA helicase Mtr4 [24-28], with orthologs now identified in the mammalian system (C1D, MPH6 and MTR4/MTREX) [1]. Structural studies of the budding yeast and mammalian RNA exosome reveal that Rrp6/EXOSC10, Rrp47/C1D and Mpp6/MPH6 interact with the complex through conserved interfaces that form composite sites for interactions with other cofactors such as Mtr4/MTR4/MTREX [22, 29-32]. The Mtr4 helicase assists in RNA substrate unwinding and plays a critical role in RNA exosome processing of the 5.8S rRNA precursor (7S rRNA) [24, 33]. Mtr4 also acts as part of larger complexes that aid the RNA exosome in nuclear RNA quality control, including the budding yeast TRAMP (Trf4/5-Air1/2-Mtr4 Polyadenylation) complex and the mammalian NEXT (Nuclear Exosome Targeting) complex [11, 34-39]. Several studies that have dissected the role of Mtr4 in aiding the RNA exosome were performed in the *Saccharomyces cerevisiae* system, establishing a number of *mtr4* mutations that disrupt specific interactions and functions of the helicase [33, 34, 40-42]. Thus, genetic model systems are a tractable system to investigate interactions with these nuclear cofactors that impact RNA exosome function and studies in such systems can expand our understanding of the influence the RNA exosome can exert over various cellular processes and pathways [43].

Given the variety of RNA exosome target RNAs and their link to many cellular processes, connections between the RNA exosome and human disease are not surprising. Many different human disease-linked mutations have been identified in genes encoding RNA exosome subunits

[12]. Mutations in *DIS3*, which encodes the catalytic component of the RNA exosome in humans [44], are the fourth most common single nucleotide variation identified in multiple myeloma (~10% of all newly diagnosed patients)[45, 46]. Multiple myeloma, which is a currently incurable cancer of the long-lived antibody-secreting plasma cells of the bone marrow, is the second most common hematologic malignancy accounting for 10-15% of incidence and 20% of deaths related to cancer of the blood and bone marrow [47, 48]. Multiple myeloma-associated *DIS3* mutations disrupt proper RNA degradation and processing in both mammalian cells and budding yeast mutant cells [49-51]. However additional mechanistic studies are required to understand how mutations in *DIS3*, and the function of the RNA exosome, could contribute to pathogenesis in multiple myeloma.

Human disease mutations have also been identified in the genes encoding the non-catalytic, structural subunits of the RNA exosome. Clinical studies have linked mutations in *EXOSC* genes to various, tissue-specific human pathologies comprising a growing family of diseases termed “RNA exosomopathies” [52-58]. RNA exosomopathy mutations have been found in all three genes that encode the cap subunits (*EXOSC1/2/3*) and several ring subunit genes (*EXOSC5/8/9*), with most being missense mutations that result in single amino acid substitutions in highly conserved domains of the subunits. Most RNA exosomopathy diseases are neurological, with mutations in *EXOSC1*, *EXOSC3*, *EXOSC5*, *EXOSC8*, and *EXOSC9* causing forms of cerebellar atrophy/degeneration and neuronopathies [52-54, 57-59]. In contrast, patients with RNA exosomopathy mutations in *EXOSC2* have a complex syndrome known as SHRF that is characterized by short stature, hearing loss, retinitis pigmentosa and distinctive facies (OMIM #617763) [55]. *In vivo* studies characterizing some of these *EXOSC* RNA exosomopathy mutations in *Saccharomyces cerevisiae* and *Drosophila melanogaster* suggest these pathogenic substitutions

could differentially impact the function of the RNA exosome complex potentially through changes in RNA targeting and cofactor interactions [53, 60-65]. Modeling these pathogenic amino acid substitutions in the budding yeast RNA exosome is an invaluable tool as several RNA exosomopathies have a small patient population, making analysis with patient tissue samples challenging. Therefore, by utilizing the budding yeast system, we can begin elucidating the functional and molecular consequences resulting from human disease mutations in RNA exosome genes [43].

In this study, we identify and characterize missense mutations in genes that encode the structural subunits of the human RNA exosome within multiple myeloma patients. We surveyed the ongoing longitudinal Multiple Myeloma Research Foundation (MMRF) study “Relating Clinical Outcomes in Multiple Myeloma to Personal Assessment of Genetic Profile” (CoMMpass) [ClinicalTrials.gov Identifier NCT01454297] to identify mutations in structural RNA exosome genes within multiple myeloma patients [66]. We focused on characterizing *EXOSC2* M40T, a missense mutation that encodes an amino acid substitution EXOSC2 p.Met40Thr (M40T) in a highly conserved region of this cap subunit that interacts with the RNA helicase MTR4. To assess the effects of this amino acid substitution in EXOSC2 on the function of the RNA exosome, we utilized the budding yeast model system and generated a variant of the *S. cerevisiae* EXOSC2 ortholog, Rrp4, which models the patient EXOSC2 M40T substitution, Rrp4 M68T. As a comparative control within our studies, we included the Rrp4 G226D variant which models a SHRF-linked pathogenic amino acid substitution in EXOSC2 p.Gly198Asp [64]. The *rrp4-G226D* cells, corresponding to the SHRF *EXOSC2* exosomopathy mutation, have defects in RNA exosome function, and are the only other budding yeast model of a disease-linked *EXOSC2* mutation [64]. Our results show that the *rrp4-M68T* gene variant can replace the function of the essential *RRP4*

gene. The *rrp4-M68T* and *rrp4-G226D* mutants show similar increases in specific RNA exosome target transcripts, suggesting shared defects in RNA processing. However, the *rrp4-M68T* mutant exhibits distinct negative genetic interactions with RNA exosome cofactor mutants, particularly *mtr4* mutants. A binding assay provides evidence that the M68T substitution impairs the interaction of Rrp4 with Mtr4. Combined, our results suggest that the Rrp4 M68T amino acid substitution, which models the multiple myeloma associated substitution EXOSC2 M40T, alters RNA exosome function by impacting the essential interaction between the complex and Mtr4. These data are the first *in vivo* characterization of this isolated multiple myeloma-associated mutation and give insight into the critical and conserved interactions between the RNA exosome and its cofactors.

3.3 Materials and Methods

Media and Chemicals

All media were prepared by standard procedures [67]. Unless stated otherwise, all chemicals were acquired from Fisher Scientific (Pittsburgh, PA), Sigma-Aldrich (St. Louis, MO), or United States Biological (Swampscott, MA).

In silico protein structure predictions

The mCSM-PPI2 platform [68], and the PyMOL viewer (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) (PyMOL) were used for structural modeling. Platforms were used with the cryo-EM structure (PDB 6D6Q) of the human nuclear RNA exosome at 3.45Å resolution [22] and the X-Ray diffraction structure (PDB 6FSZ) of the budding yeast nuclear RNA exosome at 4.60Å [21]. The ConSurf server [69-71] was used to assess the evolutionary conservation of the structure of both EXOSC2 and Rrp4.

***Saccharomyces cerevisiae* strains and plasmids**

All DNA manipulations were performed according to standard procedures [72]. *S. cerevisiae* strains and plasmids used in this study are listed in Table S1. The *rrp4Δ* (yAV1103), *rrp4Δ mpp6Δ* (ACY2471), and *rrp4Δ rrp47Δ* (ACY2474) strains have been previously described [64, 73, 74]. The *RRP45-TAP* (ACY2789) strain was obtained from Horizon Discovery Biosciences Limited and was previously described [75]. The *mtr4Δ* (ACY2532) strain was constructed by deletion of the genomic *MTR4* ORF in a wild-type (W303) strain harboring a [*MTR4*; *RRP4*;

URA3] (pAC3714) maintenance plasmid by homologous recombination using *MTR4-UTR natMX4*. This *mtr4Δ* (ACY2532) strain was then used for consecutive deletion of the genomic *RRP4* ORF to generate the *rrp4Δmtr4Δ* (ACY2536) strain as previously described [64].

Construction of the untagged *RRP4* and *rrp4-G226D* plasmids (pAC3656 and pAC3659) and the 2x-Myc tagged *RRP4* and *rrp4-G226D* plasmids (pAC3669 and pAC3672) that contain native 3' UTRs was reported previously [64]. The *rrp4-M68T LEU2 CEN6* (pAC4206) and *rrp4-M68T-2xMyc LEU2 CEN6* (pAC4207) plasmids were generated by site-directed mutagenesis of the *RRP4* (pAC3656) or *RRP4-2xMyc* (pAC3669) plasmids using oligonucleotides containing the M68T missense mutation (Fwd

5'GAAAATACGTACCGTGACCTCTCGTCCAGATAGGGTCATCAGTGACC 3', Rev

5'GGTCACTGATGACCCTATCTGGACGAGAGGTCACGGTACGTATTTTC 3') and the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The *mtr4-F7A-F10A* (pAC4099) plasmid was generated as described previously [64]. Similarly, the other *mtr4* mutant plasmids were constructed by site-directed mutagenesis of the *MTR4 HIS CEN6* plasmid (pAC4096) with the QuikChange II Site-Directed Mutagenesis Kit (Agilent) and oligonucleotides containing the corresponding missense mutations. The oligonucleotides used to generate the *mtr4-I* plasmid (pAC4103) contain the C942Y missense mutation (Fwd

5'CAAGCAGCAGCATTATTATCATACTTTGCATTCCAAGAACGCTG 3', Rev

5'CAGCGTTCTTGGAATGCAAAGTATGATAATAATGCTGCTGCTTG 3'). The oligonucleotides used to generate the *mtr4-R349E-N352E* plasmid (pAC4100) contain the R349E and N352E missense mutations [36] (Fwd 5'

GGTTGACGAAAAAAGTACCTTCGAAGAGGAAGAATTCCAAAAAGCAATGGCGTCC

3', Rev 5'

GGACGCCATTGCTTTTTGGAATTCTTCCTCTTCGAAGGTACTTTTTTCGTCAACC 3’);

The oligonucleotides used to generate the *mtr4-R1030A* plasmid (pAC4104) contain the R1030A missense mutation [33] (Fwd

5’CGTTGATCAGAATGTTCAAGGCATTAGAGGAATTGGTGAAGG 3’, Rev

5’CCTTCACCAATTCCTCTAATGCCTTGAACATTCTGATCAACG 3’) and the

oligonucleotides used to generate the *mtr4-E1033W* plasmid (pAC4105) contain the E1033W missense mutation [33] (Fwd

5’GAATGTTCAAGAGATTAGAGTGGTTGGTGAAGGAGCTGGTAGAC 3’, Rev

5’GTCTACCAGCTCCTTCACCAACCCTCTAATCTCTTGAACATTC). Plasmids were confirmed through DNA sequencing.

***Saccharomyces cerevisiae* transformations and growth assays**

All *S. cerevisiae* transformations were conducted following the standard Lithium Acetate (LiOAc) protocol [76]. Strains were grown in liquid YEPD (1% yeast extract, 2% peptone, 2% dextrose, in distilled water) in a rotating shaker at 30°C overnight to saturation. Cultures were normalized to a concentration of OD600 = 0.33 in 10 mL YEPD, then incubated at 30°C for 3-8 hours depending on the severity of their growth defect. Cells were washed and resuspended to a concentration of 2×10^9 cells/mL using TE/LiOAc. Single-stranded carrier DNA (5 μ L; 10 mg/mL), PEG/TE/LiOAc (300 μ L), and depending on reaction purpose, desired PCR product DNA or plasmid DNA, were added to cells. The mixture was incubated at 30°C in a shaker for 30 minutes. Following this incubation, DMSO (35 μ L) was added and the cells were heat shocked for 15 minutes at 42°C, washed, and plated onto selective media.

Standard plasmid shuffle assays were performed to assess the *in vivo* function of the *rrp4* variants as well as genetic interaction with RNA exosome cofactor mutants. The *rrp4* Δ (yAV1103) cells containing a *RRP4 URA3* maintenance plasmid and transformed with vector (pRS315) and transformed with *RRP4* (pAC3656), *rrp4-G226D* (pAC3659), *rrp4-M68T* (pAC4206), *RRP4-2xMyc* (pAC3669) or *rrp4-M68T-2xMyc* (pAC4207) plasmid were grown on Ura⁻ Leu⁻ minimal media control plates, which select for cells that contain both the *RRP4 URA3* maintenance plasmid as well as the *RRP4/rrp4 LEU2* plasmid, and 5-FOA Leu⁻ minimal media plates, which select for cells that lack the *RRP4 URA3* maintenance plasmid and contain only the *RRP4/rrp4 LEU2* plasmid. The plates were incubated at 30°C for 2-3 days and single colonies from the 5-FOA Leu⁻ minimal media plates were selected in quadruplicate and streaked onto selective Leu⁻ minimal media plates. The cells containing only the *RRP4/rrp4 LEU2* plasmid are referred to as *RRP4*, *rrp4-G226D* or *rrp4-M68T* cells. A similar strategy was used to generate *mtr4* Δ (ACY2532) cells that contain only the *MTR4* (pAC4096) or *mtr4-1* (pAC4103) *HIS3*, *CEN6* plasmid. The *mtr4* Δ cells transformed with *MTR4* or *mtr4-1* were grown overnight and serially diluted and spotted onto Ura⁻ His⁻ minimal media plates and 5-FOA minimal media plates, which select for cells that lack the *URA3* maintenance plasmid and contain only the *MTR4/mtr4 HIS3* plasmid. Single colonies of cells containing only *MTR4* or *mtr4-1 HIS3* plasmid were collected in quadruplicate and are referred to as *MTR4* or as *mtr4-1* cells.

The *in vivo* function of the *rrp4-M68T* variant was assessed in growth assays on solid media and in liquid culture. For growth on solid media, *rrp4* Δ (yAV1103) cells containing only *RRP4* (pAC3656), *rrp4-G226D* (pAC3659) or *rrp4-M68T* (pAC4206) were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation. Cell concentrations were normalized to an OD₆₀₀ = 1, and samples were serially diluted in 10-fold dilutions and spotted onto Leu⁻ minimal media

plates, Leu⁻ minimal media plates containing 25 μ M fluorouracil (5-FU), YEPD plates or YEPD plates containing 3% formamide, 150 mM hydroxyurea or 5 μ g/ml phleomycin. Plates were grown at 25°C, 30°C, and 37°C for 2-3 days. For growth in liquid culture, cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation, diluted to an OD₆₀₀ = 0.05 in Leu⁻ minimal media in a 24-well plate, and growth at 37°C was monitored and recorded at OD₆₀₀ in a BioTek® SynergyMx microplate reader with Gen5™ v2.04 software over 36 hrs. For these liquid growth assays, the cells incubate in the microplate reader for many hours before their density is within the dynamic range of the machine to record the doubling times. For the results shown, each sample was performed in at least 3 independent biological replicates with 3 technical replicates for each biological sample. Doubling times were calculated using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com), GraphPad Software, San Diego, California USA.

Immunoblotting

To analyze protein expression levels of C-terminally Myc-tagged Rrp4 and Rrp4 M68T, *rrp4* Δ (yAV1103) cells expressing only Rrp4-2xMyc (pAC3669) or Rrp4-M68T-2xMyc (pAC4207) were incubated in 2 mL of Leu⁻ minimal medium at 30°C and grown to saturation overnight. The 10 mL cultures with an OD₆₀₀ = 0.2 were prepared and incubated at 30°C or 37°C for 5 hr. Yeast cell pellets were collected by centrifugation and transferred to 2 mL screw-cap tubes. Cell pellets were flash frozen with liquid nitrogen and stored at -20°C. Yeast cell lysate was prepared by resuspending pellets in 0.3 mL of RIPA-2 Buffer (50 mM Tris-HCl, pH 8; 150 mM NaCl; 0.5% sodium deoxycholate; 1% NP40; 0.1% SDS) supplemented with protease inhibitors [1 mM PMSF; 3 ng/ml PLAC (pepstatin A, leupeptin, aprotinin, and chymostatin)],

followed by addition of 300 μ l glass beads. Lysates were placed in a Mini Bead Beater 16 Cell Disrupter (Biospec) for 6×1 min at 25°C with ice submersion intervals of 1 minute between rounds, and then centrifuged at 4°C at 12,000 RPM for 10 min. Protein lysate concentration was determined by Pierce BCA Protein Assay Kit (Life Technologies). Whole cell lysate protein samples (40 μ g) in reducing sample buffer (50 mM Tris HCl, pH 6.8; 100 mM DTT; 2% SDS; 0.1% Bromophenol Blue; 10% Glycerol) were resolved on Criterion 4–20% gradient denaturing gels (Bio-Rad), transferred to nitrocellulose membranes (Bio-Rad) and Myc-tagged Rrp4 proteins were detected with anti-Myc monoclonal antibody 9B11 (1:2000; Cell Signaling). The 3-phosphoglycerate kinase (Pgk1) protein was detected using anti-Pgk1 monoclonal antibody (1:30,000; Invitrogen) as a loading control. For quantitation, ImageJ v1.4 software (National Institutes of Health, MD; <http://rsb.info.nih.gov/ij/>) was used to measure protein band areas and intensities. Protein percentages relative to Pgk1 were calculated using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com), GraphPad Software, San Diego, California USA.

Co-Immunoprecipitations

To assess association of Rrp4 M68T with the RNA exosome complex, we utilized *RRP45-TAP* (ACY2789) cells expressing *RRP4-2xMyc* (pAC3669), *rrp4-G226D-2xMyc* (pAC3672) or *rrp4-M68T-2xMyc* (pAC4207) and immunoprecipitated Rrp45-TAP using the IgG Sepharose beads as previously described [64]. Briefly, cell samples were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation and 10-20 mL cultures with an OD₆₀₀ = 0.2 were prepared and grown at 30°C for 5 hr. Yeast cell lysates were prepared by resuspending cell pellets in 0.5-0.75 mL of IPP150 Buffer (10mM Tris-HCl, pH 8; 150 mM NaCl; 0.1% NP40, PMSF) supplemented with protease inhibitors [1 mM PMSF; Pierce™ Protease Inhibitors

(Thermo Fisher Scientific)], and 300 μ L of glass beads. Cells were disrupted in a Mini Bead Beater 16 Cell Disrupter (Biospec) for 4-5 \times 1 min at 25°C with 1 min on ice between repetitions. Crude lysate was transferred to a chilled microcentrifuge tube and remaining beads were washed with an additional 150 μ L of IPP150 Buffer. Lysate was then cleared by centrifugation at 16,000 \times g for 10 min at 4°C. Protein lysate concentration was determined by Pierce BCA Protein Assay Kit (Life Technologies). For input samples, 40 μ g of cleared lysate was collected and frozen at -20°C. For co-immunoprecipitations, 1 mg of cleared lysate in IPP150 Buffer was prepared, 30 μ L of a 1:1 bead slurry of IgG Sepharose® 6 Fast Flow Beads (GE Healthcare) was added, and samples were incubated at 4°C overnight with mixing. After overnight incubation, beads were washed three times in 1 mL IPP150 Buffer for 5 min each (IgG Sepharose beads). Whole cell lysate input samples (40 μ g) and total bound samples in reducing sample buffer were boiled for 5 min[®] at 100°C, resolved on 4–20% Criterion™ TGX precast polyacrylamide gels (Bio-Rad), transferred to nitrocellulose membranes (Bio-Rad). Levels of associated Rrp4-Myc proteins with the Rrp45-TAP tagged subunit were detected by immunoblotting. Myc-tagged Rrp4 proteins were detected with mouse anti-Myc monoclonal antibody 9B11 (1:2000; Cell Signaling). TAP-tagged Rrp45 protein was detected with peroxidase anti-peroxidase (PAP) soluble complex antibody produced in rabbit (1:5000, Sigma-Aldrich). The 3-phosphoglycerate kinase (Pgk1) protein was detected using anti-Pgk1 monoclonal antibody (1:30,000; Invitrogen) as a loading control.

To assess the association of Mtr4 with Rrp4 M68T, we utilized $\Delta rrp4$ cells expressing *RRP4-2xMyc* (pAC3669), or *rrp4-M68T-2xMyc* (pAC4207). These cells were transformed with either an empty plasmid (*pAC1*) or a plasmid exogenously 2x-FLAG tagged Mtr4 (pAC3719). Briefly, cell samples were grown in 2 mL His⁻Leu⁻ minimal media overnight at 30°C to

saturation and 40 mL cultures with an $OD_{600} = 0.2$ were prepared and grown at 30°C for 4 hr. Yeast cell lysates were prepared as described above. Cleared protein lysate concentration was determined by Pierce BCA Protein Assay Kit (Life Technologies). For input samples, 50 μg of cleared lysate was collected and frozen at -20°C . For co-immunoprecipitations, 1.75 mg of cleared lysate in IPP150 Buffer was prepared, 15 μL of a 1:1 bead slurry of PierceTM Anti-c-Myc Magnetic Beads (ThermoFisher) was added, and samples were incubated at 4°C overnight with mixing. After overnight incubation, beads were washed three times in 1 mL IPP150 Buffer for 15 sec each. Whole cell lysate input samples (50 μg) and total bound samples in reducing sample buffer were boiled for 10 min[®] at 100°C , resolved on 4–20% CriterionTM TGX precast polyacrylamide gels (Bio-Rad), transferred to nitrocellulose membranes (Bio-Rad). Levels of associated Mtr4-FLAG proteins with Rrp4-Myc or Rrp4 M68T-Myc tagged subunit were detected by immunoblotting. Myc-tagged Rrp4 proteins were detected with mouse anti-Myc monoclonal antibody 9B11 (1:2000; Cell Signaling). FLAG-tagged Mtr4 protein was detected with anti-Myc monoclonal antibody 9A3 (1:1000; Cell Signaling). The 3-phosphoglycerate kinase (Pgk1) protein was detected using anti-Pgk1 monoclonal antibody (1:30,000; Invitrogen) as a loading control. A peroxidase AffiniPure Goat Anti-Mouse IgG, light chain specific secondary was used for detecting the bound proteins <70 kDa. Quantitation was performed first by standardizing detected FLAG-tagged and Myc-tagged protein levels to detected Pgk1 levels. Then protein levels from bound samples were normalized to unbound levels to generate a fraction bound.

Genetic Interaction Analysis

To test genetic interactions between *rrp4-M68T* and RNA exosome cofactor/subunit deletion mutants, *rrp4Δ mpp6Δ* (ACY2471), and *rrp4Δ rrp47Δ* (ACY2474) cells containing only *RRP4* (pAC3656), *rrp4-G226D* (pAC3659) or *rrp4-M68T* (pAC4206) were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation, serially diluted, and spotted on Leu⁻ minimal media plates. The plates were incubated at 30°C or 37°C for 3 days. Cells were also grown in liquid culture as described in *S. cerevisiae* transformation and growth assays method. The *rrp4Δ mpp6Δ* (ACY2471) cells containing only *RRP4* (pAC3656), *rrp4-G226D* (pAC3659) or *rrp4-M68T* (pAC4206) were further assayed by being serially spotted onto Leu⁻ minimal media plates containing 25 μM 5-FU, YEPD plates or YEPD plates containing 3% formamide.

To test for genetic interactions between *rrp4-M68T* and *mtr4* mutants, *mtr4-F7A-F10A*, *mtr4-I*, *mtr4-R1030A*, and *mtr4-E1033W*, *rrp4Δ mtr4Δ* (ACY2536) cells containing the [*MTR4*; *RRP4*; *URA3*] (pAC3714) maintenance plasmid were transformed with *RRP4* (pAC3656), *rrp4-G226D* (pAC3659) or *rrp4-M68T* (pAC4206) *LEU2* plasmid and selected on Ura⁻Leu⁻ minimal media plates. Transformed cells containing both the *URA3* maintenance plasmid and the *RRP4/rrp4* variant plasmid were subsequently transformed with *MTR4* (pAC4096), *mtr4-F7A-F10A* (pAC4099), *mtr4-I* (pAC4103), *mtr4-R1030A* (pAC4104), or *mtr4-E1033W* (pAC4105) *HIS3* plasmid and selected on Ura⁻Leu⁻His⁻ minimal media plates. The transformed cells were then streaked to onto 5-FOA Leu⁻ His⁻ plates to select for cells that did not contain the *URA3* maintenance plasmid. The resulting *rrp4Δ mtr4Δ* cells containing only *RRP4*, *rrp4-G226D* or *rrp4-M68T* *LEU2* plasmid and *MTR4* or *mtr4* variant *HIS3* plasmid were grown in 2 mL Leu⁻ His⁻ minimal media overnight at 30°C to saturation, serially diluted, and spotted on Leu⁻ His⁻ minimal media plates. The plates were incubated at 30°C and 37°C for 3 days. Cell growth was quantified on a scale from 0 to 5 across triplicate assays, with a score of “0” representing

lethality and a score of “5” representing full growth across dilutions. Scores were averaged and displayed as a heatmap using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com), GraphPad Software, San Diego, California USA.

Total RNA isolation

Total RNA from *RRP4*, *rrp4-G226D*, *rrp4-M68T*, *MTR4* or *mtr4-1* cells was isolated using the MasterPure™ Yeast RNA Purification Kit (Epicentre, Lucigen). Cells were incubated in 2 mL of Leu- minimal medium at 30°C and grown to saturation overnight. Cultures were diluted in 10 mL to an $OD_{600} = 0.2$ and further incubated at 37°C for 5 hours. Cells were pelleted by centrifugation, transferred to RNase-free microcentrifuge tubes and flash frozen with liquid nitrogen. Frozen cell pellets were stored at -80°C. RNA isolation was performed according to the MasterPure™ Yeast RNA Purification Kit (Epicentre, Lucigen) manufacturer’s protocol. Total RNA was resuspended in 50 μ L DEPC-treated water and stored at -80°C.

RT-qPCR

All oligonucleotides used in this study are shown in Table S2. For analysis of steady-state RNA levels using quantitative PCR, three independent biological replicates of *RRP4*, *rrp4-G226D*, *rrp4-M68T*, *mtr4-1* and *MTR4* cells were grown in 2 mL Leu⁻ or His⁻ minimal media overnight at 30°C. Cultures (10 mL) with an $OD_{600} = 0.2$ were prepared from the saturated

cultures and cells were grown at 37°C for 5 hr. Total RNA was isolated from cell pellets as described and 1 µg of total RNA was reverse transcribed to first strand cDNA using the M-MLV Reverse Transcriptase (Invitrogen) according to manufacturer's protocol. Quantitative PCR was performed on technical triplicates of cDNA (10 ng) from three independent biological replicates using gene specific primers (0.5 mM; Table S2), QuantiTect SYBR Green PCR master mix (Qiagen) on a StepOnePlus Real-Time PCR machine (Applied Biosystems; Tanneal=55°C, 44 cycles). *ALG9* or *PGK1* was used as an internal control. The mean RNA levels were calculated by the $\Delta\Delta C_t$ method [77]. Statistical analysis comparing the control cells (*RRP4* or *MTR4*) and the mutant cells (*rrp4* or *mtr4-1*) was performed by t-test ($\alpha < 0.05$) using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com), GraphPad Software, San Diego, California USA.

DATA AVAILABILITY STATEMENT

Strains and plasmids summarized in Table S1 and Table S2 are available upon request. Genomic data from CoMMpass are available at dbGaP with the accession number phs000748.v7.p4. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article, figures, and tables.

3.4 Results

EXOSC2 p.Met40Thr substitution is located within a conserved region of the cap subunit that interacts with the essential helicase MTR4.

Mutations in the gene *DIS3*, which encodes the catalytic component of the RNA exosome, are commonly found in patients diagnosed with multiple myeloma [45, 46], suggesting a link between RNA exosome function and disease pathology. We therefore considered whether mutations in the other components of the RNA exosome would be found in multiple myeloma patients. Missense mutations in *EXOSC* genes, which encode the structural subunits of the RNA exosome, were identified in multiple myeloma patients through interrogating the ongoing longitudinal Multiple Myeloma Research Foundation (MMRF) study “Relating Clinical Outcomes in Multiple Myeloma to Personal Assessment of Genetic Profile” (CoMMpass) [ClinicalTrials.gov Identifier NCT01454297]. A total of 1,154 newly diagnosed multiple myeloma patients were enrolled in CoMMpass and profiled by genomic testing and tissue sampling throughout treatment. The molecular profiling collected through CoMMpass reveals several rare missense mutations within *EXOSC* genes (Figure S1). One patient missense mutation identified within exon 1 of *EXOSC2* encodes EXOSC2 p.Met40Thr (M40T), which is located in a highly conserved region of the N-terminal domain of EXOSC2 (Figure 1C). Notably, EXOSC2 Met40 lies within a key binding interface between the human RNA exosome and the RNA helicase MTR4 [22].

The patient with the *EXOSC2 M40T* mutation also had chromosomal aberrations including a chromosomal translocation t(11;14) and hyperdiploidy disease. The chromosomal translocation t(11;14) is an IgH translocation which is an initiating event that occurs frequently in multiple myeloma (~15-20% of patients) [78]. From the CoMMpass dataset, we determined that the variant allele frequency is 0.2266, however the copy number of the chromosome 9 *EXOSC2* locus is 2.6, suggesting that this *EXOSC2* allele is found on the extra copy of ch9 that is present in over half the cells in the patient. Based on these findings, we conclude that this hyperdiploidy of

chromosome 9 occurred after the t(11;14) translocation event and that the *EXOSC2* mutation either co-occurred with the chromosomal gain or shortly after.

To explore how *EXOSC2* M40T could alter the function of the RNA exosome complex, we modeled the *EXOSC2* M40T amino acid substitution using a recent structure of the human RNA exosome in complex with the essential RNA helicase MTR4 [22]. MTR4 makes several direct contacts with the RNA exosome, forming a binding interface with a total surface area of $1,440\text{\AA}^2$ [22]. Among the direct contacts between MTR4 and the RNA exosome complex, the N-terminal domain of *EXOSC2* interacts with MTR4 through an aliphatic surface that includes Met40. As shown in Figure 2A, *EXOSC2* Met40 engages with a hydrophobic pocket of MTR4 including I1014. An amino acid substitution of Thr40, while unlikely to disrupt the aliphatic surface, could disrupt the hydrophobic interaction at this contact given the polar, shortened side chain of threonine (Figure 2A). The *EXOSC2* M40T substitution could therefore destabilize the interface between the N-terminal domain of *EXOSC2* and MTR4. We also modeled the amino acid substitution in the budding yeast *EXOSC2* ortholog, Rrp4, using a recent structure of the *S. cerevisiae* RNA exosome [32]. As shown in Figure 2B, the budding yeast RNA exosome in complex with Mtr4 shows structural similarities to the human complex, with Rrp4 interacting directly with Mtr4. Rrp4 Met68, corresponding to the *EXOSC2* Met40 residue, engages with the helicase directly through hydrophobic interactions at the binding interface. This binding interface of yeast Mtr4 is part of a large hydrophobic pocket which includes F924 and I923 (not labeled), similar to hydrophobic pocket in the human MTR4 helicase. Introduction of Thr40 would most likely disrupt this contract, similar to our predictions for the M40T substitution in *EXOSC2*. Furthermore, the region surrounding Rrp4 Met68 is structurally synonymous to the aliphatic

surface surrounding EXOSC2 Met40, allowing for us to assess the effects of the EXOSC2 M40T amino acid substitution at the conserved interface within the yeast system.

We further explored the conservation of the binding interface of human EXOSC2 as compared to budding yeast Rrp4 using the bioinformatics tool ConSurf (Figure S2). The ConSurf server estimates the evolutionary conservation of amino acids of a protein based on phylogenetic trees between homologous sequences, providing conservation rates for each residue that reflect both functional and structural importance [69-71]. Consistent with the sequence alignment (Figure 1C) and structural modeling, this tool predicts high conservation for both EXOSC2 Met40 and Rrp4 Met68 and surrounding residues (Figure S2). Additionally, ConSurf estimates high conservation rates at each site of contact between EXOSC2 and Rrp4 with the helicase MTR/Mtr4, further supporting the evolutionary importance of this interaction.

***Saccharomyces cerevisiae rrp4-M68T* mutant cells that model the multiple myeloma *EXOSC2 M40T* variant show sensitivity on drugs that impact RNA metabolism.**

To assess the functional consequences of the EXOSC2 M40T amino acid substitution, we generated the corresponding amino change in the *S. cerevisiae* ortholog Rrp4, M68T (Figure 1C). We first performed a plasmid shuffle growth assay in which cells deleted for the genomic copy of *RRP4* are transformed with plasmids containing different *rrp4* alleles (See *Materials and Methods*). This approach ensures that the background for all variants that are compared to one another is identical [79]. This growth assay reveals the *rrp4-M68T* allele can replace the essential *RRP4* gene as the *rrp4-M68T* cells grow similarly to control cells expressing wild-type *RRP4* at all temperatures examined (Figure 3A). As a comparison, we included cells expressing the *rrp4-*

G226D allele as the sole copy of the essential *RRP4* gene. The *rrp4-G226D* mutant allele models a known SHRF pathogenic amino acid change that has been shown to cause impaired RNA exosome function *in vivo* [64]. As previously reported, cells expressing only *rrp4-G226D* show impaired growth at 37°C (Figure 3A). Furthermore, we assessed the growth of both the *rrp4-M68T* and *rrp4-G226D* mutant cells using a liquid growth assay and quantified doubling times (Figure 3B, 3C). These data confirm that the growth of *rrp4-M68T* cells does not differ significantly from wild-type *RRP4* cells.

To explore whether the *rrp4-M68T* mutation sensitizes cells to altered RNA processing, we tested for growth defects when cells are grown on media containing 5-fluorouracil (5-FU) [80, 81] (Figure 3D). The *rrp4-M68T* cells show a slight growth defect compared to wild-type *RRP4* cells at 30°C when grown on solid media containing 25 µM 5-FU. This growth defect is more evident when the cells are challenged with both 37°C and 25 µM 5-FU. As a comparison, the *rrp4-G226D* cells show a severe growth defect when grown on solid media containing 5-FU both at 30°C and 37°C. To further explore whether the *rrp4-M68T* cells exhibit other changes in cell growth, we tested for growth defects when cells are grown on media containing chemicals that disrupt different cellular pathways (Figure 3E). Formamide alters RNA metabolism [82], hydroxyurea impairs DNA synthesis [83], and phleomycin acts as a mutagen by introducing double-strand breaks in DNA [84]. The *rrp4-M68T* cells do not show any increased sensitivity when grown at 30°C on solid media containing 3% formamide, 150 mM hydroxyurea or 5 µg/ml phleomycin (Figure 3E). In contrast, the *rrp4-G226D* cells show enhanced growth defects at 30°C when grown on solid media containing 3% formamide, 150 mM hydroxyurea or 5 µg/ml phleomycin. Taken together, these data suggest that the *rrp4-M68T* cells are sensitive to defects in RNA processing but do not exhibit the same extent of disrupted cellular pathways as the

previously studied *rrp4-G226D* cells, which model a pathogenic RNA exosomopathy mutation that has severely impaired RNA exosome function *in vivo* [64].

***rrp4-M68T* cells have impaired RNA exosome function in processing RNA targets linked to Mtr4-RNA exosome interactions.**

To further assess the *in vivo* consequences on RNA exosome function of the *rrp4-M68T* variant, we measured the steady-state levels of several RNA exosome targets in *rrp4-M68T* cells using RT-qPCR. We assessed the steady-state level of precursor RNAs that are targeted by the RNA exosome and are impacted by *mtr4* mutant alleles, including the telomerase component RNA *TLC1*, which is processed by the RNA exosome in a manner dependent on TRAMP complex association, and the 3' extended forms of *U4* snRNA and *snR33* snoRNA [10, 85, 86]. In this analysis, we included both *rrp4-G226D* and *mtr4-1* cells as comparative controls. The *mtr4-1* cells have a missense mutation in *MTR4* that results in accumulation of polyadenylated targets within the nucleus [34, 40-42]. We detect increases in the steady-state level of both mature and precursor *TLC1* in *rrp4-M68T* cells similar to that observed in *rrp4-G226D* cells (Figure 4A) [64]. Both mature and precursor *TLC1* steady-state levels are significantly increased in *mtr4-1* cells. Furthermore, we detect a significant increase in the steady-state level of the 3' extended forms of the *U4* snRNA and *snR33* snoRNA in the *rrp4-M68T* cells (Figure 4B and 4C). This increase in the levels of the extended form of these target RNAs is also observed in the *rrp4-G226D* cells and, to an even larger extent, the *mtr4-1* cells. We also assessed steady-state levels of 5.8S rRNA precursors in *rrp4-M68T* and found no accumulation compared to wild-type *RRP4* cells (Figure S3), although levels of both mature 5.8S rRNA and pre-5.8S rRNA do increase in *rrp4-G226D*

cells which supports previous observations that the *rrp4-G226D* cells exhibit accumulation of 7S rRNA [64].

We also measured the steady-state level of select targets that are impacted within *rrp4-G226D* cells [64]. We assessed the target *INO1*, which encodes inositol-3-phosphate synthetase [87, 88]. *INO1* mRNA has previously been identified as a transcript bound to the catalytic subunit of the RNA exosome [89], and was the most significantly decreased transcript in a previous RNA-Seq analysis of the *rrp4-G226D* cells [64]. In *rrp4-M68T* cells, the steady-state level of *INO1* is significantly decreased, similar to results for *rrp4-G226D* (Figure 4D). We also assessed three cryptic unstable transcripts (CUTs) that accumulate within *rrp4-G226D* cells [64]. The steady-state levels of these three CUTs are not significantly increased in *rrp4-M68T* cells (Figure 4E). Taken together, these data suggest that the *rrp4-M68T* cells have some molecular consequences resulting from the modeled multiple myeloma amino acid substitution; however, they differ from those resulting from the modeled SHRF substitution in the *rrp4-G226D* cells.

The Rrp4 M68T variant can associate with the RNA exosome complex.

The sensitivity of the *rrp4-M68T* cells to drugs that impact RNA processing (Figure 3D) and the observed accumulation of key RNA exosome target RNAs (Figure 4) suggest that RNA exosome function may be impaired by the modeled multiple myeloma amino acid substitution. Previous studies suggest that SHRF-linked amino acid substitutions modeled in Rrp4 affect the RNA exosome function in part by disrupting complex integrity [64]. To assess the impact of the modeled multiple myeloma amino acid substitution on the association of Rrp4 with other RNA exosome core subunits, we first assayed the protein level of Rrp4 M68T. We measured the steady-

state level of the Myc-tagged Rrp4 M68T subunit when expressed as the sole copy of the Rrp4 protein in *rrp4Δ* cells grown at either 30°C or 37°C (Figure 5A). Immunoblotting reveals that the steady-state level of Rrp4 M68T is comparable to wild-type Rrp4 at both temperatures tested. We next performed co-immunoprecipitations using *RRP45-TAP* cells that contain the endogenous, genomic *RRP4* gene and express a C-terminally tandem affinity purification (TAP)-tagged Rrp45 core subunit from the endogenous *RRP45* locus. We expressed Rrp4-Myc or Rrp4 M68T-Myc from plasmids in these *RRP45-TAP* cells. The Rrp45-TAP protein was immunoprecipitated and association of the Myc-tagged Rrp4 variants was assayed by immunoblotting (Figure 5B). Under these conditions in which an endogenous copy of *RRP4* is present, we can detect association of Rrp4 M68T-Myc with Rrp45-TAP at levels equal to that of Rrp4-Myc. As a comparative control, we also performed co-immunoprecipitations with *RRP45-TAP* cells expressing an exogenous Rrp4 G226D-Myc variant. Under these conditions with an endogenous copy of *RRP4* present, we cannot detect association of Rrp4 G226D-Myc with the TAP-tagged core subunit, supporting previous observations [64]. Taken together, these data suggest that Rrp4 M68T is biochemically similar to a wild-type Rrp4 subunit, and the multiple myeloma amino acid substitution likely has no impact on RNA exosome complex integrity.

The *rrp4-M68T* mutant shows negative genetic interactions with *mtr4* mutants that impact TRAMP complex association and RNA helicase unwinding.

As the Rrp4 M68T variant associates with the RNA exosome complex and has a steady level equivalent to wild-type Rrp4, the observed sensitivity to disrupted RNA processing in *rrp4-M68T* cells (Figure 3D) and accumulation of select RNA exosome target transcripts (Figure 4) could be due to altered interaction between Rrp4 and Mtr4. As depicted in Figure 6A, the nuclear

RNA exosome cofactors Mpp6 and Rrp47 and the associated exonuclease Rrp6 aid in recruiting Mtr4 to the RNA exosome [31]. Rrp6 and Rrp47 form a composite site that binds to the N-terminus of Mtr4, recruiting the helicase to the RNA exosome [29]. Mtr4 forms contacts with the cap subunit Rrp4 and the cofactor Mpp6, stabilizing the helicase on the RNA exosome complex through a very conserved interface between the cap subunit and the helicase [22, 30]. The interaction between Mtr4 and Rrp4 provides a surface for the RNA exosome to associate with the TRAMP complex, which helps facilitates nuclear RNA surveillance and quality control of ncRNA [1, 23, 25, 26, 35, 37-39, 90-92]. In addition to Mtr4, the TRAMP complex is composed of a zinc-knuckle RNA binding protein, Air1 or Air2, and a non-canonical oligo(A) polymerase, Trf4 or Trf5, that oligoadenylates RNA [93]. The TRAMP complex triggers degradation by adding short polyadenylated tails to the 3' end of substrate RNA and delivering them to the RNA exosome [23, 86, 93, 94].

To assess genetic interactions between Mtr4 and the RNA exosome in *S. cerevisiae* modeling the multiple myeloma patient mutation, we performed an analysis of a series of five *mtr4* mutant alleles that introduce amino acid substitutions in Mtr4 as summarized in Figure 6B. We also included the *rrp4* mutant variant, *rrp4-G226D*, for comparison as this *rrp4* variant has a negative interaction with the *mtr4-F7A-F10A* mutant allele [64]. Included within our five *mtr4* alleles is the *mtr4-1* allele which is a known temperature sensitive mutant [34, 40-42]. These genetic data are shown both as representative solid growth assays (Figure 6C) and as a heatmap (Figure 6D), which summarizes data from three independent replicates for all these genetic experiments. As predicted, *RRP4 mtr4-1* cells have a severe growth defect at 37°C that is shared by both double mutant *rrp4-M68T mtr4-1* and *rrp4-G226D mtr4-1*.

Two of the five *mtr4* mutant alleles, *mtr4-F7A-F10A* and *mtr4-R349E-N352E*, impair protein-protein interactions of Mtr4 in *S. cerevisiae*. The Mtr4 F7A F10A variant disrupts Mtr4 interactions with Rrp6/Rrp47 by introducing two amino acid substitutions, F7A and F10A, into the N-terminus of Mtr4 (Figure 6B) [29]. The *mtr4-R349E-N352E* mutant allele impairs the association of Mtr4 with the poly(A) RNA polymerase Trf4 within the TRAMP complex and thus disrupts the recruitment of the TRAMP complex to the RNA exosome [36]. The *rrp4-M68T mtr4-F7A-F10A* double mutant cells grow similarly to *rrp4-M68T* cells at 30°C, however at 37°C the double mutant cells show a mild growth defect in comparison to the single mutant *rrp4-M68T* and the *RRP4 mtr4-F7A-F10A* cells (Figure 6D). As shown previously, the *rrp4-G226D mtr4-F7A-F10A* show a severe growth defect at 37°C compared to *rrp4-G226D* cells [64]. The *rrp4 M68T mtr4-R349E-N352E* double mutant cells show severe growth defects at 30°C and 37°C and compared to the single mutant *rrp4-M68T* cells or the *RRP4 mtr4-R349E-N352E* control cells. Contrastingly, the *rrp4 G226D mtr4-R349E-N352E* double mutant cells show no difference in growth at 30°C compared to the *RRP4 mtr4-R349E-N352E* control cells and improved growth compared to the single mutant *rrp4-G226D* cells at 37°C (Figure 6C and 6D).

The final two *mtr4* mutant alleles that we tested for genetic interaction with the *rrp4-M68T* variant impact nucleic acid unwinding by Mtr4 [33]. Studies of an RNA-bound Mtr4 structure demonstrate that residues of R1030 and E1033 mediate key nucleic acid base interactions with the helicase helical bundle [33, 34]. Mutagenesis of these residues in *S. cerevisiae*, generating the mutant alleles *mtr4-R1030A* and *mtr4-E1033A*, reveal that these residues play important but distinct roles in helicase activity [33]. The *rrp4-M68T mtr4-R1030A* double mutant cells are not viable at either temperature tested. In contrast the *rrp4-M68T mtr4-E1033A* double mutant cells are viable and grow similar to the *RRP4 mtr4-E1033A* control cells at both 30°C and 37°C. The

growth defect of *rrp4-G226D* at 37°C is too severe to assess genetic interactions with either the *mtr4-R1030A* or the *mtr4-E1033A* mutation under these growth conditions. However, in contrast to the lethality observed for the *rrp4-M68T mtr4-R1030A* double mutant, the *rrp4-G226D mtr4-R1030A* and *rrp4-G226D mtr4-E1033A* double mutant cells have comparable growth to the single mutant *rrp4-G226D* cells as well as the *RRP4 mtr4-R1030A* and *RRP4 mtr4-E1033A* control cells at 30°C. The *rrp4-M68T* double mutants that show synthetic lethality are viable when rescued by expression of a wild-type *RRP4* plasmid (Figure S4), demonstrating that the growth defects and lethality observed are due to negative genetic interactions between the *rrp4* and *mtr4* mutants. Taken together, these data suggest that the *rrp4-M68T* cells have negative genetic interactions with specific *mtr4* mutant alleles, distinct from those previously described for the *rrp4-G226D* mutant model.

The *rrp4-M68T* mutant shows negative genetic interactions with *mpp6Δ*.

As depicted in Figure 6A, the nuclear RNA exosome cofactors Mpp6 and Rrp47 and the exoribonuclease Rrp6 help to recruit and stabilize the interaction with Mtr4. The exosome cofactor Rrp47 interacts with and stabilizes the exoribonuclease Rrp6 and the cofactor Mpp6 interacts with the nuclear RNA exosome through direct contacts with the cap subunit Rrp40 [19, 29, 31]. To further evaluate the impact that the modeled multiple myeloma amino acid substitution may have on the RNA exosome-Mtr4 interaction *in vivo*, we tested whether the *rrp4-M68T* variant exhibits genetic interactions with *mpp6* or *rrp47* mutants by deleting these non-essential, nuclear exosome cofactor genes *MPP6* and *RRP47* in combination with *rrp4-M68T*. For comparison, we included the *rrp4-G226D* variant as these cells have known negative genetic interactions with these mutants [64]. We examined the growth of these double mutants relative to single mutants (*rrp4-M68T* or

rrp4-G226D) and the control mutant cells (*RRP4 mpp6Δ* or *RRP4 rrp47Δ*) in solid and liquid media growth assays (Figure 7). In the solid media growth assays the *rrp4-M68T mpp6Δ* double mutant cells show growth very similar to the *rrp4-M68T* cells at both 30°C and 37°C after both one and two days of growth (Figure 7B). The *rrp4-M68T rrp47Δ* cells show a severe growth defect at 37°C compared to the single mutant *rrp4-M68T*; however, this impaired growth is comparable to that of the *RRP4 rrp47Δ* cells, which has been previously reported for the single mutant *rrp47Δ* [27] (Figure 6C). In contrast, the *rrp4-G226D mpp6Δ* and *rrp4-G226D rrp47Δ* double mutant cells show a severe growth defect at both temperatures compared to the single mutant *rrp4-G226D* cells as described previously [64].

While the solid media growth assay suggests comparable growth between the controls and the *rrp4-M68T* double mutant cells, the liquid media growth assay reveals a modest growth defect of the *rrp4-M68T mpp6Δ* at 37°C compared to both the *rrp4-M68T* and control *RRP4 mpp6Δ* cells (Figure 7D), with the doubling time significantly longer than that of wild-type *RRP4* cells (Figure 7E). The liquid growth assay also shows doubling times for *rrp4-M68T rrp47Δ* and *RRP4 rrp47Δ* double mutants are nearly twice that of wild-type *RRP4* cells, but do not differ significantly when compared to each other (Figure 7D and 7E). The observed growth defect of the *rrp4-M68T mpp6Δ* double mutant in liquid culture is revealed in a solid media growth assay when the cells are challenged with formamide or 5-FU (Figure 7F). The distinct growth defect of the *rrp4-G226D mpp6Δ* double mutant is also exacerbated by growth on these chemicals. Taken together, these data suggest a negative genetic interaction between the *rrp4* variants and *mpp6* mutants, with both double mutants having exacerbated defects when challenged with drugs that impact RNA metabolism.

The essential helicase Mtr4 has less association with the Rrp4 M68T variant compared to wild-type Rrp4.

Given the negative genetic interactions between *rrp4-M68T* and both *mpp6* mutants and *mtr4* mutants and the interaction between both EXOSC2 Met40 and Rrp4 Met68 and the RNA helicase Mtr4 [22, 32], we predicted that the interaction between the RNA exosome and Mtr4 would be affected in *rrp4-M68T* cells. Previous studies investigated the physical interaction between the Rrp4 G226D and Mtr4 using co-immunoprecipitations and the data suggested that there is decreased association between the RNA exosome and the helicase in *rrp4-G226D* cells [64]. We employed a similar approach to investigate whether the physical interaction between Mtr4 and the RNA exosome is impacted by the Rrp4 M68T variant. We performed a co-immunoprecipitation with cells expressing Rrp4-Myc or Rrp4 M68T-Myc as the sole copy of Rrp4 and co-expressing Mtr4-FLAG (Figure 8). The Rrp4-Myc proteins were immunoprecipitated and association with Mtr4-FLAG was assayed by immunoblotting. As shown in Figure 8, there is a significant decrease in the amount of Mtr4-FLAG that co-immunoprecipitates with Rrp4 M68T-Myc as compared to Rrp4-Myc (Figure 8A-B). The amount of Rrp4 M68T-Myc and Rrp4-Myc is comparable (Figure 8C), showing that the difference in detected Mtr4-FLAG is not due to decreased protein levels or inefficient immunoprecipitation of Rrp4 M68T-Myc (Figure 8C). These data, therefore, suggest that Mtr4 association with the Rrp4 cap subunit is significantly disrupted by the Rrp4 M68T amino acid substitution. Combined with the genetic data (Figure 6 and Figure 7), the structural modeling (Figure 2), and the increased steady-state level of RNA exosome target RNAs (Figure 4), these results strongly suggest that there is destabilization of the interaction between Mtr4 and the RNA exosome complex in *rrp4-M68T* cells that impacts the function of the molecular machine.

3.5 Discussion

In this study, we modeled and analyzed a multiple myeloma patient *EXOSC2* mutation in the *S. cerevisiae* homolog *RRP4*. We generated *rrp4-M68T* mutant cells expressing the variant Rrp4 M68T, which corresponds to the *EXOSC2* M40T variant. Analysis of these *rrp4-M68T* cells reveals that this amino acid substitution affects RNA exosome function. While our biochemical assays show that the Rrp4 M68T variant can associate with the RNA exosome complex and function as the sole copy of the essential Rrp4 RNA exosome cap subunit, *rrp4-M68T* cells do show growth defects when grown in media containing drugs that impact RNA processing. The *rrp4-M68T* cells also show accumulation of known RNA exosome targets. These defects in RNA exosome function could result from an impaired interaction between the complex and the essential RNA helicase Mtr4 as predicted by structural modeling. Our genetic analyses support this model as *rrp4-M68T* cells show negative genetic interactions with both *mpp6* and *mtr4* mutants. Furthermore, we demonstrate through a co-immunoprecipitation assay that the M68T substitution in Rrp4 decreases association with the Mtr4 helicase. These data suggest that the introduction of the multiple myeloma associated amino acid change could impact the binding interface between *EXOSC2* and *MTR4*, potentially impairing the function of the essential RNA exosome *in vivo* for a subset of Mtr4-dependent targets.

Structural studies reveal the evolutionary conservation of the interaction between the RNA exosome and Mtr4 (Figure S2), with the helicase cofactor interacting with the complex through multiple points of contact including a direct interface with *EXOSC2*/Rrp4 and indirect stabilizing interactions with the cofactors Mpp6, Rrp47 and the associated exonuclease Rrp6 [22, 30]. This robust interaction between the complex and the essential helicase likely explains why the *rrp4-M68T* cells show no functional consequences unless challenged through introduction of drugs

impacting RNA processing or loss of other stabilizing cofactors, such as in *rrp4-M68T mpp6Δ* double mutant cells. While this model would also predict a negative genetic interaction between *rrp47Δ* and *rrp4-M68T*, the *rrp4-M68T rrp47Δ* cells show a growth defect at 37°C that is indistinguishable from that of *RRP4 rrp47Δ* cells. This growth defect in *rrp4-M68T rrp47Δ* and *RRP4 rrp47Δ* cells is likely due to the loss of Rrp6 association with the RNA exosome complex given the stabilizing role Rrp47 plays for Rrp6 [19, 27]. The growth defects resulting from destabilization of Mtr4 in *rrp4-M68T rrp47Δ* cells is likely masked by the larger consequence of disassociating Rrp6 from the complex. We do detect a slight growth defect in *rrp4-M68T* cells expressing an *mtr4* variant that disrupts the stabilizing interactions between Rrp6, Rrp47 and Mtr4 (*mtr4-F7A-F10A*), pointing to the importance of the Rrp4-Mtr4 interface. We do, however, observe significant molecular consequences in the *rrp4-M68T* cells. We detect accumulation of several documented RNA exosome target transcripts, particularly those linked to RNA exosome-Mtr4 association [10, 86]. Through biochemical assays we also observe a significant decrease in interaction between Mtr4 and Rrp4 M68T as compared to wild-type Rrp4. This decreased association further suggests that the modeled multiple myeloma mutation destabilizes the interaction between the RNA exosome and the essential RNA helicase. Taken together, these data suggest that while the consequences resulting *in vivo* from the Rrp4 M68T variant are subtle at the macro scale, they are impactful molecularly for a specific set of target RNAs and for the biochemical interaction between the RNA exosome and Mtr4.

The interaction between the RNA exosome and Mtr4 could also be critical for other interactions, particularly those involving the TRAMP (Trf4/5-Air1/2-Mtr4 Polyadenylation) complex. Our genetic analyses reveal a negative genetic interaction between *rrp4-M68T* and *mtr4-R349E-N352E*. The Mtr4 R349E N352E variant impairs Mtr4-Trf4 binding and impacts TRAMP

complex assembly *in vivo* [36]. The *rrp4-M68T* cells that express Mtr4 R349E N352E as the sole copy of the helicase grow very poorly at both 30°C and 37°C as compared to control *RRP4 mtr4-R349E-N352E* cells, suggesting TRAMP complex assembly and association with the RNA exosome may also be impacted by Rrp4 M68T. Intriguingly, we also detect synthetic lethality for the *rrp4-M68T mtr4-R1030A* double mutant. This lethality is specific to *rrp4-M68T mtr4-R1030A* cells as the *rrp4-M68T* cells expressing the other helicase mutant, *mtr4-E1033W*, show growth similar to the control (*RRP4 mtr4-E1033W*). Both Mtr4 R1030A and Mtr4 E1033W decrease helicase unwinding capability [33]. However, the Mtr4 R1030A variant is also implicated in disrupting target discrimination by the TRAMP complex, potentially by disrupting preferential polyadenylation by Trf4 [33]. Therefore, the negative genetic interaction observed for *rrp4-M68T rrp4-R1030A* cells further suggests that TRAMP function is impacted in *rrp4-M68T* cells. Taken together with our structural modeling data, we hypothesize that a stabilized interaction between the RNA exosome and Mtr4 is necessary for TRAMP association and the slightest perturbation, even a subtle destabilization at one contact point with the helicase, could disrupt this vital interaction between TRAMP and the complex. More biochemical studies could be performed to explore how changes within the EXOSC2-Mtr4 interface impact the interaction with the TRAMP complex.

Our studies also show that *rrp4-M68T* mutant cells have distinct genetic interactions as compared to the *rrp4-G226D* cells. The *rrp4-G226D mtr4-R349E-N352E* double mutant cells surprisingly show improved growth at 37°C compared to either single mutant. Even more surprising is the synthetic lethality in cells expressing *rrp4-G226D* and either *mtr4* helicase mutant (*mtr4-R1030A* and *mtr4-E1033W*). These genetic interactions could suggest that the modeled SHRF amino acid substitution (Rrp4 G226D) has distinct *in vivo* consequences compared to the

modeled multiple myeloma-associated substitution Rrp4 M68T. The Rrp4 G226D variant has decreased association with Mtr4 and the *rrp4-G226D* cells show transcriptomic differences from wild-type cells consistent with disrupted RNA exosome-Mtr4 interactions [64]. Similarly, we observe decreased association between the Rrp4 M68T variant and Mtr4 and some RNA exosome target transcripts accumulate in *rrp4-M68T* cells that also accumulate in *rrp4-G226D* cells. However, notably, we do not detect any changes in select CUTs or 5.8S rRNA precursors in *rrp4-M68T* cells. We do, intriguingly, observe a significant decrease in the steady state level of *INO1* mRNA in *rrp4-M68T* cells that is shared in the *rrp4-G226D* cells [64]. Previous work characterizing the *rrp4-G226D* mutation suggested that the significant change in *INO1* mRNA levels could reflect defects in the cytoplasmic roles of the RNA exosome [64], though the exact molecular mechanism remains unknown. A comparison of the results obtained for *rrp4-M68T* and *rrp4-G226D* suggests that there may be some distinct defects in RNA exosome function in each of these mutants though they may also have some overlapping consequences *in vivo* in part due to altered association with Mtr4. The difference in molecular consequences between the two *rrp4* variants could be attributed to the impact on RNA exosome complex integrity observed in *rrp4-G226D* cells, which was not observed in *rrp4-M68T* cells (Figure 5B) [64].

The difference in severity of functional and molecular consequences we observe for the *rrp4-M68T* and *rrp4-G226D* mutant models may partially explain the differences in disease pathology between SHRF patients with the mutation *EXOSC2 G198D* and the multiple myeloma patient with the mutation *EXOSC2 M40T*. The *EXOSC2 G198D* mutation was identified in SHRF patients through whole exome sequencing and classified as causing a novel Mendelian syndrome [55]. In contrast, the *EXOSC2 M40T* mutation is a spontaneous, somatic mutation that likely co-occurred with a chromosome 9 duplication. Additionally, the patient with this *EXOSC2 M40T*

mutation has several chromosomal aberrations that are a hallmark of multiple myeloma, suggesting these *EXOSC2* mutations could be passenger mutations rather than a pathogenic driver of the multiple myeloma. Upon further analysis of the noncoding mutations in the patient harboring *EXOSC2 M40T*, we found a second mutation present in intron 1 of *EXOSC2* in this patient (Figure S6). This mutation (*EXOSC2 SNV chr9:130,693,915 T>G*) is predicted to alter the splice donor site and likely result in a misprocessed mRNA or truncated protein. Through RNA-Seq data available in CoMMpass for this patient, we determined that the *EXOSC2 M40T* missense mutation and the splice donor mutation are expressed from the same allele. Interestingly, we calculate the allelic frequency of these two *EXOSC2* mutations to be very similar (0.2266 vs. 0.2191). This suggests that *EXOSC2 M40T* and the *EXOSC2* splice donor mutation either co-occurred or that the splice donor mutation was selected for in response to the *EXOSC2 M40T* missense mutation, which could negatively affect cell growth and/or survival. As *EXOSC2* is an essential gene in 1076/1086 cancer cell lines in the Cancer Dependencies Map project (depmap.org) including all 19 myeloma cell lines in the dataset, a future approach would be to CRISPR mutate the *EXOSC2 M40T* mutation within myeloma cell lines to determine the effects on exosome function and myeloma cell growth and survival.

As the *rrp4-M68T* cells show defects in function of the RNA exosome likely through altered of interactions with the RNA helicase Mtr4 and the associated TRAMP complex, this *EXOSC2 M40T* substitution could be detrimental to the function of the human RNA exosome. Altering key cofactor interactions with the RNA exosome could impact the processing and degradation of target RNA transcripts such as small ncRNA species that have key regulatory roles in various cellular processes. Furthermore, the interaction between the RNA exosome and MTR4 has been suggested to resolve secondary DNA structures associated with strand asymmetric DNA

mutagenesis that can lead to genome instability and chromosomal translocations particularly in plasma B cells [95]. The high level of evolutionary conservation within the N-terminus of EXOSC2 that interacts with MTR4 (Figure 1C and Figure S2), suggests that there could be evolutionary pressure to maintain the integrity of certain sequences within EXOSC2 that specifically interact with key cofactors. Taking a genetic approach to assess different *EXOSC* missense mutations associated with human diseases can help unravel different consequences in specific interactions of the essential RNA exosome complex.

Utilizing the yeast genetic model system, we have characterized an *EXOSC2* mutation found in a multiple myeloma patient. However, this mutation was one of several mutations identified in genes encoding structural subunits of the RNA exosome in the CoMMpass study (Figure S1). The frequency of multiple myeloma mutations identified in the *DIS3* catalytic exosome gene suggests that there is an important link between multiple myeloma and the RNA exosome. By modeling identified *EXOSC* mutations in the budding yeast system, we can examine whether these mutations impair the function of the essential RNA exosome and provide a deeper understanding of the role this conserved complex may have in cancer pathologies. While it is unlikely that the identified *EXOSC* mutations drive the multiple myeloma disease, our study here clearly shows that these mutations have *in vivo* consequences for the conserved and essential RNA exosome-Mtr4 interaction. In addition, by studying other models of *EXOSC* disease linked mutations, such as those identified in RNA exosomopathy patients, we can provide insight into the biological pathways that are altered in these different disorders. As more pathogenic mutations are uncovered in *EXOSC* genes through patient genomic screenings, generating *in vivo* models to explore the consequences of these changes can help to define the most critical interactions of the

complex with various cofactors thus expand our understanding of the biological functions of this singular, essential RNA processing and degradation complex.

3.6 Chapter III Acknowledgements

We thank members of the Corbett lab for critical discussions and input. We thank Dr. Benjamin Barwick for his contributions in assisting with the analysis of CoMMpass dataset. We also thank Dr. Graeme Conn and Dr. Ambro van Hoof for their intellectual contributions. This work was supported by National Institutes of Health (NIH) R01 grants (GM058728) and the NIH- funded Emory Initiative for Maximizing Student Development (R25 GM099644) to A.H.C and National Cancer Institute (NCI) R21 grant (CA273773) to L.H.B. M.C.S. was supported by a National Institute of General Medical Sciences (NIGMS) F31 grant (GM134649-01). S.E.S. was supported by the National Science Foundation (NSF) Graduate Research Fellowship (GRFP 1937971). Lastly, we would also like to thank the Genetics Society of America and the [Saccharomyces Genomic Database](#) (SGD) [99] for providing community resources and support for scientific discovery.

3.7 Chapter III Figures

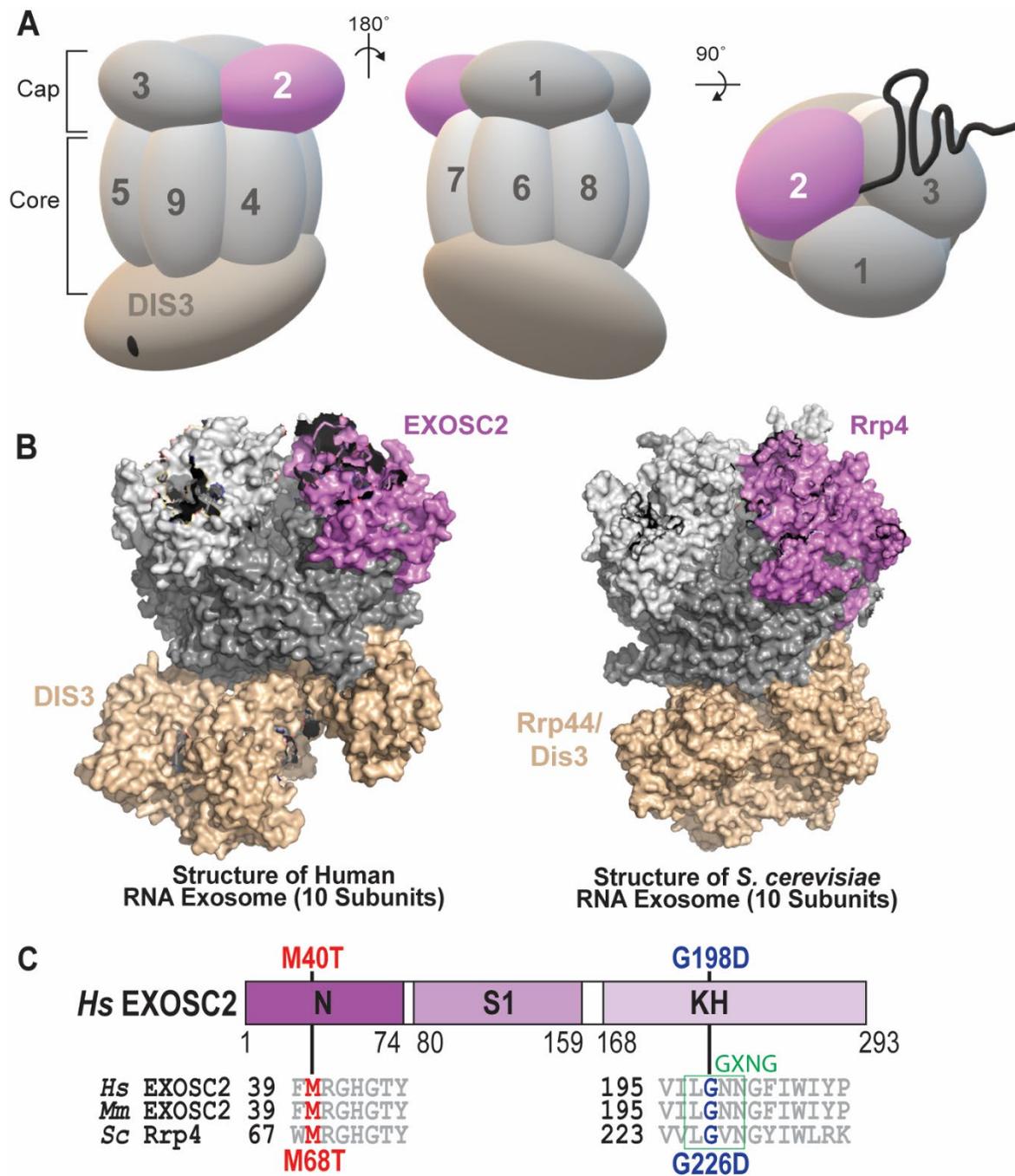


Figure 1. Overview of multiple myeloma linked amino acid substitutions in the human cap subunit EXOSC2 of the RNA exosome.

(A) The RNA exosome is an evolutionary conserved ribonuclease complex composed of nine structural subunits (EXOSC1-9) and one catalytic subunit (DIS3) that form a “Cap” and “Core” ring-like structure. The 3-subunit cap at the top of the complex is composed of EXOSC1/Cs14 (Human/*S. cerevisiae*), EXOSC2/Rrp4, and EXOSC3/Rrp40 (labeled 1-3). The 6-subunit Core is composed of EXOSC4/Rrp41, EXOSC5/Rrp46, EXOSC6/Mtr3, EXOSC7/Rrp42, EXOSC8/Rrp43, and EXOSC9/Rrp45 (labeled 4-9). The DIS3/Dis3/Rrp44 catalytic subunit is located at the bottom of the complex. Together the cap and core form a barrel-like structure through which RNA is threaded to the catalytic DIS3/Dis3/Rrp44 subunit. Recent missense mutation in the gene encoding the EXOSC2 cap subunit (pink) have been identified in patients presenting with multiple myeloma. (B) Structural models of the human nuclear RNA exosome (left) [PDB 6D6R] [22] and the *S. cerevisiae* nuclear RNA exosome (right) [PDB 6FSZ] [32] are depicted with the cap subunit EXOSC2/Rrp4 labeled and colored in pink. (C) Domain structure of EXOSC2/Rrp4. This cap subunit is composed of three domains: an N-terminal domain, a putative RNA binding S1 domain, and a C-terminal putative RNA binding KH (K homology) domain. A conserved “GxNG” motif identified in the KH domain is boxed in green [96]. The position of the disease-linked amino acid substitutions in human EXOSC2 are depicted above the domain structures. The amino acid substitution (p.Met40Thr) we report in a multiple myeloma patient is shown in red. An amino acid substitution (p.Gly198Asp) linked to SHRF is shown in blue [55]. Sequence alignments of EXOSC2/Rrp4 orthologs from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*) and *S. cerevisiae* (*Sc*) below the domain structures show the highly conserved residues altered in disease in red and blue and the conserved sequences flanking these residues in gray.

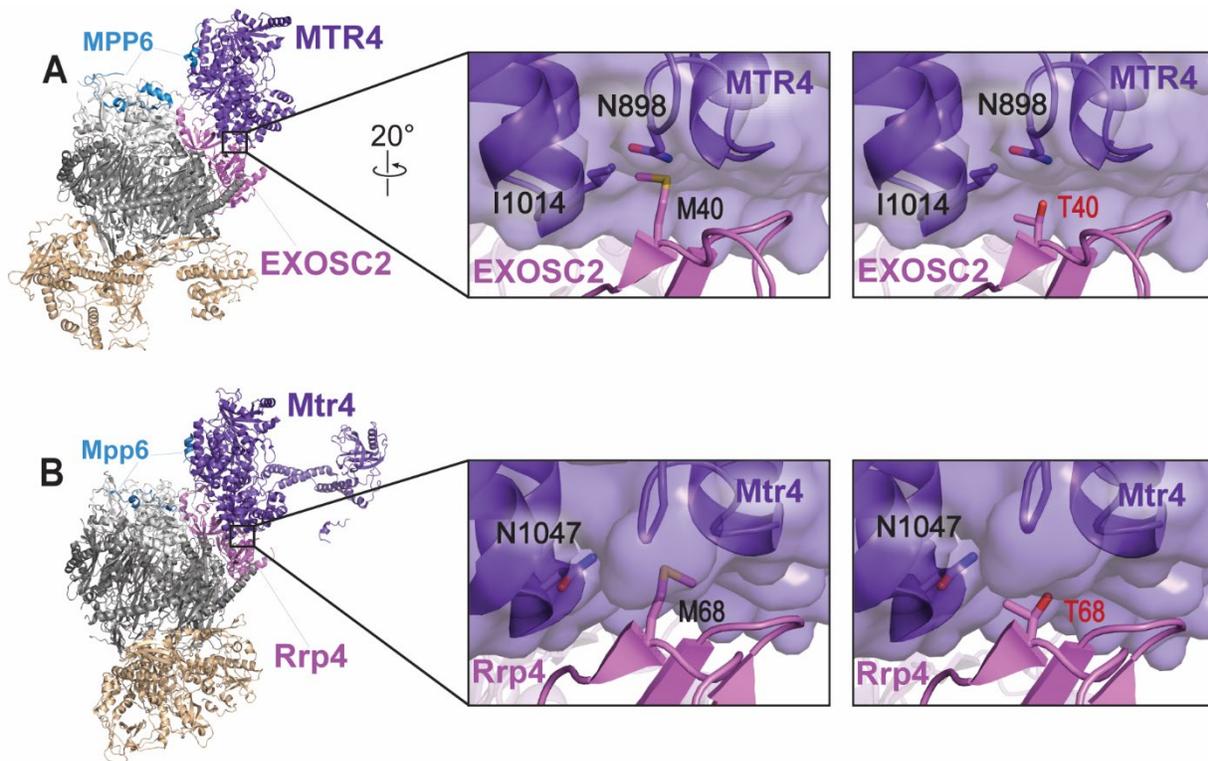


Figure 2. Modeling the multiple myeloma EXOSC2 M40T amino acid substitution in the human EXOSC2 cap subunit and the *S. cerevisiae* ortholog Rrp4.

(A) Structural modeling of the human EXOSC2 p.Met40Thr (M40T) amino acid substitution identified in a patient with multiple myeloma (PDB 6D6R) [22]. The full structure of the human RNA exosome with the associated cofactor MTR4 (purple) is depicted with a zoomed-in representation of the interface between EXOSC2 (pink) and MTR4. Modeling of the native EXOSC2 Met40 (M40) residue (left) or the multiple myeloma-associated EXOSC2 Thr40 (T40) residue (right) is shown. The EXOSC2 Met40 residue is located in the N-terminal domain of EXOSC2, within a conserved aliphatic interface with MTR4. EXOSC2 Met40 and MTR4 associate through hydrophobic interactions, which includes contacts with MTR4 Ile1014 (I1014).

(B) Structural modeling of the budding yeast Rrp4 Met68Thr (M68T) amino acid change,

corresponding to EXOSC2 p.Met40Thr, in the budding yeast RNA exosome (PDB 6FSZ) [32]. The full structure of the budding yeast RNA exosome complex with the associated MTR4 ortholog, Mtr4 (purple), is depicted on the left. A zoomed-in representation of the interface between Rrp4 (pink) and Mtr4 are shown, modeling the native Rrp4 Met68 (M68) residue (left) or the modeled multiple myeloma associated substitution Rrp4 Thr68 (T68) residue (right). The Rrp4 Met68 residue is conserved between human and yeast and is located in the N-terminal domain of Rrp4. Similar to EXOSC2 Met40, Rrp4 Met68, associates with the helicase Mtr4 through primarily hydrophobic interactions, including contacts with several glycine residues in a neighboring loop of Mtr4. Parts of the human nuclear cofactor protein MPP6/MPH6 (blue) and the budding yeast ortholog Mpp6 (blue) are also resolved in the structures shown in (A) and (B). Both MPP6/MPH6 and Mpp6 aid in stabilizing the interaction of the RNA helicase with the RNA exosome in addition to the direct interface made between EXOSC2 and MTR4 in humans or Rrp4 and Mtr4 in budding yeast [30, 31].

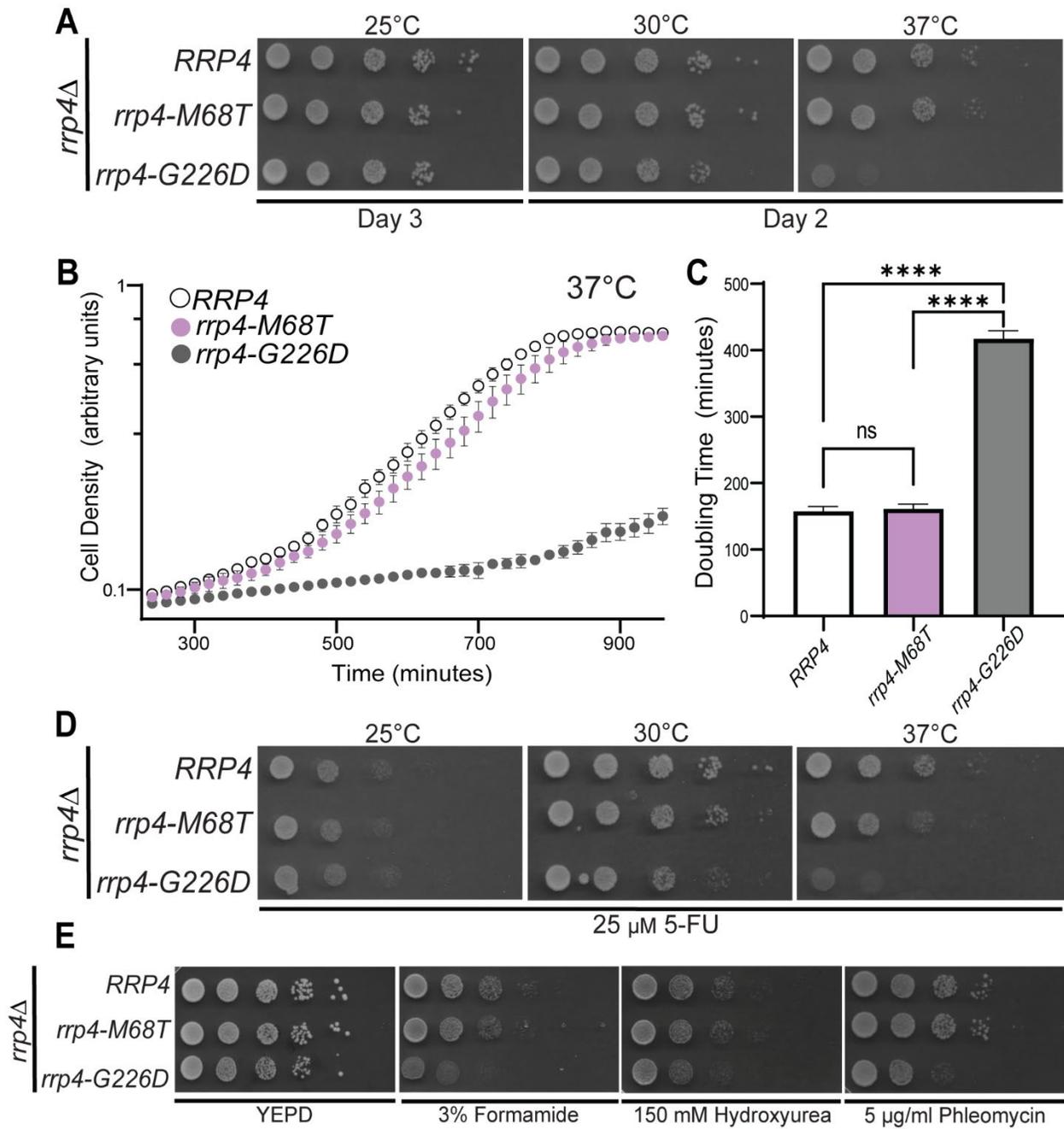


Figure 3. *S. cerevisiae rrp4-M68T* mutant cells that model the EXOSC2 M40T variant identified in multiple myeloma patients show impaired function on drugs that impact RNA processing.

S. cerevisiae cells expressing Rrp4 variants that model the multiple myeloma amino acid change or, as a control, the previously characterized [64] SHRF-linked amino acid change found in EXOSC2 were generated as described in *Materials and Methods*. (A-B) The *rrp4*Δ cells expressing only *RRP4* or mutant *rrp4* were serially diluted, spotted onto solid selective media grown at the indicated temperatures or grown in liquid media at 37°C with optical density measurement used to assess cell density over time. The doubling time of these cells grown in liquid media is quantified and graphed in (C). (D) The *rrp4*Δ cells expressing either *RRP4* or *rrp4-M68T* were serially diluted, spotted onto solid selective media containing 25 μM fluorouracil (5-FU) and grown at the indicated temperatures. Images shown are after two days of growth. (E) The *rrp4*Δ cells expressing only *RRP4* or *rrp4-M68T* were serially diluted and spotted onto solid YEPD media containing 3% formamide, 150 mM hydroxyurea or 5 μg/ml phleomycin and grown at 30°C. Images shown are after two days of growth. In all assays performed, *rrp4-G226D* cells, previously reported to be severely impaired at 37°C, were included as a control [64]. Data shown are representative of three independent experiments (n = 3).

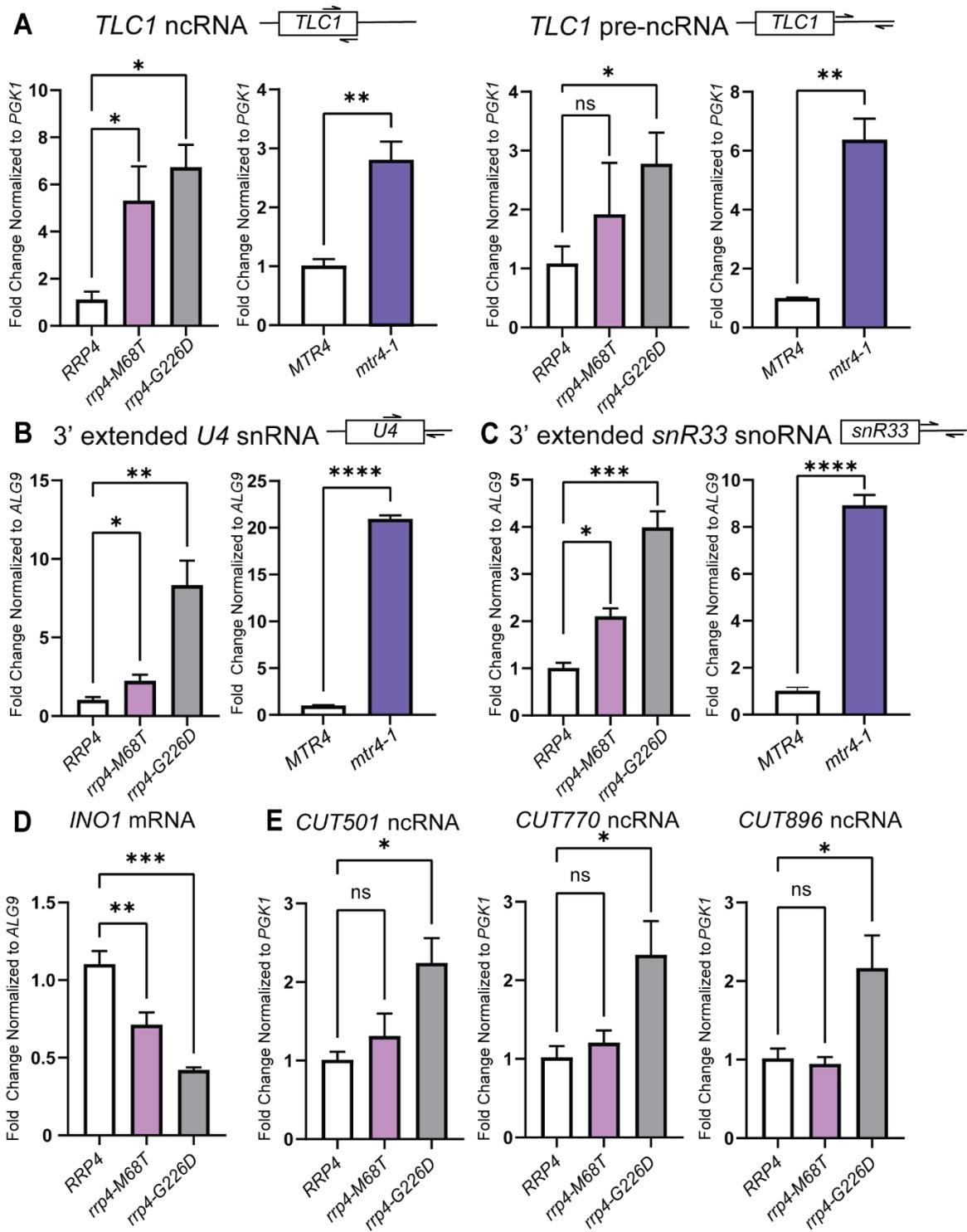


Figure 4. The *rrp4-M68T* mutant cells show elevated levels of specific RNA exosome target transcripts that depend on the Mtr4-RNA exosome interaction *in vivo*.

The steady-state level of several RNA exosome target transcripts was assessed in *rrp4-M68T* cells (denoted in pink). The steady-state levels of these RNAs were also assessed in the previously described *rrp4* variant *rrp4-G226D* as a control (denoted in gray). (A) The *rrp4-M68T* cells show an elevated steady-state level of mature *TLC1* telomerase component ncRNA relative to *RRP4* cells. The steady-state level of the precursor *TLC1* ncRNA in *rrp4-M68T* cells follows this upward trend though not statistically significant compared to *RRP4* cells. This increase in both mature and precursor *TLC1* is also observed in *mtr4-1* (denoted in purple) compared to the wild-type control *MTR4*. (B) The *rrp4-M68T* cells show an elevated steady-state level of 3'-extended pre-*U4* snRNA relative to *RRP4* cells. The *rrp4-G226D* mutant cells and *mtr4-1* cells have even higher steady-state levels of this pre-*U4* snRNA when compared to the *RRP4* and *MTR4* controls, respectively. (C) The *rrp4-M68T* cells show an elevated steady-state level of 3' extended *snR33* snoRNA relative to *RRP4* cells that is similar to the increase observed in the *rrp4-G226D* mutant cells and *mtr4-1* cells. (D) The *rrp4-M68T* cells show a decreased steady-state level of the mRNA transcript *INO1* compared to wild-type *RRP4* control cells. A decrease in this mRNA was shown previously in *rrp4-G226D* cells [64]. (E) The steady-state levels of non-coding, cryptic unstable transcripts, *CUT501*, *CUT770*, and *CUT896*, are not significantly increased in *rrp4-M68T* cells compared to control as shown previously in the *rrp4-G226D* mutant cells [64]. In (A-E), total RNA was isolated from cells grown at 37°C and transcript levels were measured by RT-qPCR using gene specific primers and graphed as described in *Materials and Methods*. Gene specific primer sequences are summarized in Table S2. The location of primers specific to the ncRNA transcripts are graphically represented by the cartoons above each bar graph. Within the cartoon transcript, the box represents

the body of the mature transcript. Error bars represent standard error of the mean from three biological replicates. Statistical significance of the RNA levels in *rrp4* variant cells relative to *RRP4* cells and in the *mtr4-1* cells relative to *MTR4* cells is denoted by an asterisk (* p -value \leq 0.05; ** p -value \leq 0.01, *** p -value \leq 0.001, **** p -value \leq 0.0001).

(top) and input (bottom) samples were analyzed by immunoblotting. As a control, immunoprecipitations were also performed from untagged *RRP45* cells (No TAP Control) expressing Myc-tagged Rrp4 and Rrp4 variants. The bound/input level of Rrp4-Myc was detected with anti-Myc antibody and bound/input level of Rrp45-TAP was detected with a peroxidase anti-peroxidase (PAP) antibody. Bound Rrp45-TAP is also detected by the anti-Myc antibody as the Protein A moiety of the TAP tag binds to the antibody. The input level of 3-phosphoglycerate kinase (Pgk1) was detected with an anti-Pgk1 antibody and shown as a loading control. Data shown is representative of three independent experiments (n = 3).

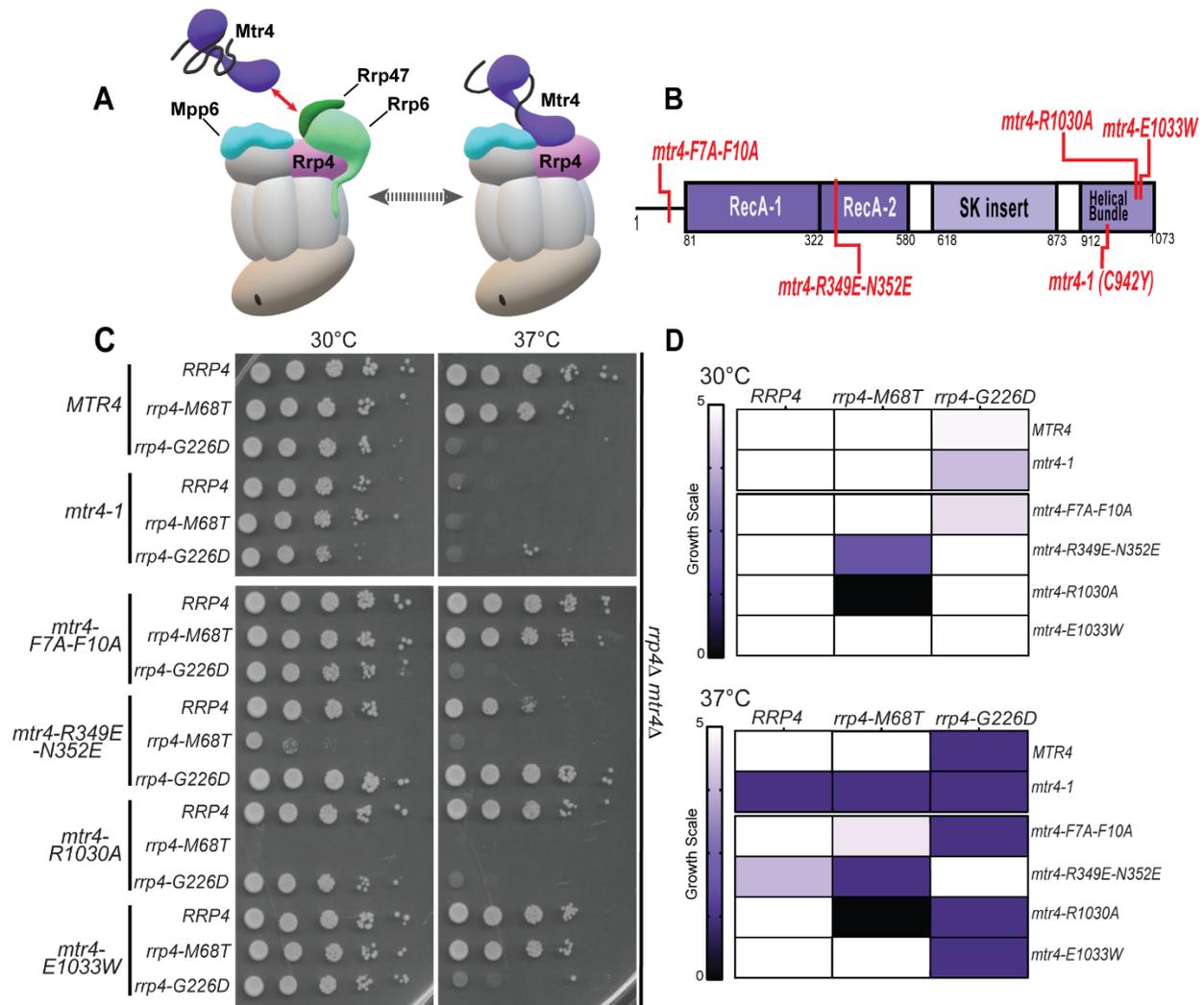


Figure 6. The *rrp4-M68T* mutant cells show specific negative genetic interactions with *mtr4* mutants that are predicted to impair the Trf4/5-Air1/2-Mtr4 (TRAMP) complex.

(A) Cartoon depicting the budding yeast nuclear RNA exosome with interacting nuclear cofactors Mpp6 (turquoise) and Rrp47 (dark green), the exoribonuclease Rrp6 (light green), and the essential RNA helicase, Mtr4 (purple) [29, 30, 32]. The association of Mtr4 with the RNA exosome is facilitated by interactions between Mtr4 and Rrp6/Rrp47 (denoted by the red arrowed line) and by interactions with Mpp6 which is associated with the Rrp40 RNA exosome subunit and the Rrp4

subunit [29, 31, 34]. The association of Mtr4 with the RNA exosome can also facilitate interaction with the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) complex, which triggers degradation of certain RNA targets by adding short oligo(A) tails to the 3' end of these targets and delivering them to the RNA exosome [23, 86, 93, 94]. In addition to Mtr4, the TRAMP complex is composed of a noncanonical poly(A) polymerase, Trf4/5, and a zinc-knuckle RNA binding protein, Air1/2 [93]. Central to the degradation of TRAMP-targeted RNAs by the RNA exosome is the association of Mtr4 with Trf4/5, Air1/2 and the cap subunits and nuclear cofactors of the RNA exosome complex [29, 36]. (B) Domain structure for *S. cerevisiae* Mtr4. The helicase has a low-complexity N-terminal sequence followed by the conserved helicase region. The helicase region is composed of two RecA domains and a helical domain (labeled helical bundle) that form the globular core typical of DExH family proteins. The helical bundle was originally described as the “ratchet” domain for its role in translocating nucleic acid by a Brownian ratchet [97]. In addition, Mtr4 contains an insertion domain and KOW domain that fold into a helical stalk (labeled SK insertion) [34, 98]. The amino acid changes used for this experiment are labeled in red along the domain structure. (C) Double mutant cells containing *rrp4-M68T* and specific *mtr4* mutants show lethality at both 30°C and 37°C. The *rrp4Δ mtr4Δ* double mutant cells were serially diluted, spotted onto solid media, and grown at the indicated temperatures for 3 days. The *mtr4* mutant plasmids included in this experiment are as follows; *mtr4-1*— a temperature sensitive mutant that contains a missense mutation resulting in the amino acid substitution Cys942Tyr, which causes accumulation of poly(A)⁺ RNA in the nucleus at 37°C [40-42]; *mtr4-F7A-F10A*—an *mtr4* allele that impairs the interaction with Rrp6/Rrp47 [29]; *mtr4-R349E-N352E*—a mutation that impairs the association of Mtr4 with the poly(A) RNA polymerase Trf4 with the Mtr4 helicase [36]; *mtr4R-1030A* and *mtr4-E1033W*—two mutations within the helical bundle that differentially

impact nucleic acid unwinding by Mtr4 [33]. *mtr4-1* mutant cells expressing *RRP4*, *rrp4-M68T* or *rrp4-G226D* show lethality at 37°C presumably due to the known temperature sensitive nature of the *mtr4-1* allele [40]. Growth of double mutant cells containing *rrp4-M68T* or *rrp4-G226D* are shown. (D) Summary of *rrp4 mtr4* mutant cell growth. Triplicate solid media assays were performed on double mutant cells containing *rrp4-M68T* or *rrp4-G226D* and the series of *mtr4* variants. Cell growth at both 30°C and 37°C was semi-quantified on a scale of 0 (lethal; black) to 5 (comparable to *RRP4* wild-type growth; white). Growth scale of the double mutant cells is represented through the color gradient on the two heatmaps. All double mutant cells were generated as described in *Materials and Methods*. Images shown are from a singular solid media growth assay with all samples plated on the same -Leu media plate. Data is representative of three independent experiments (n = 3).

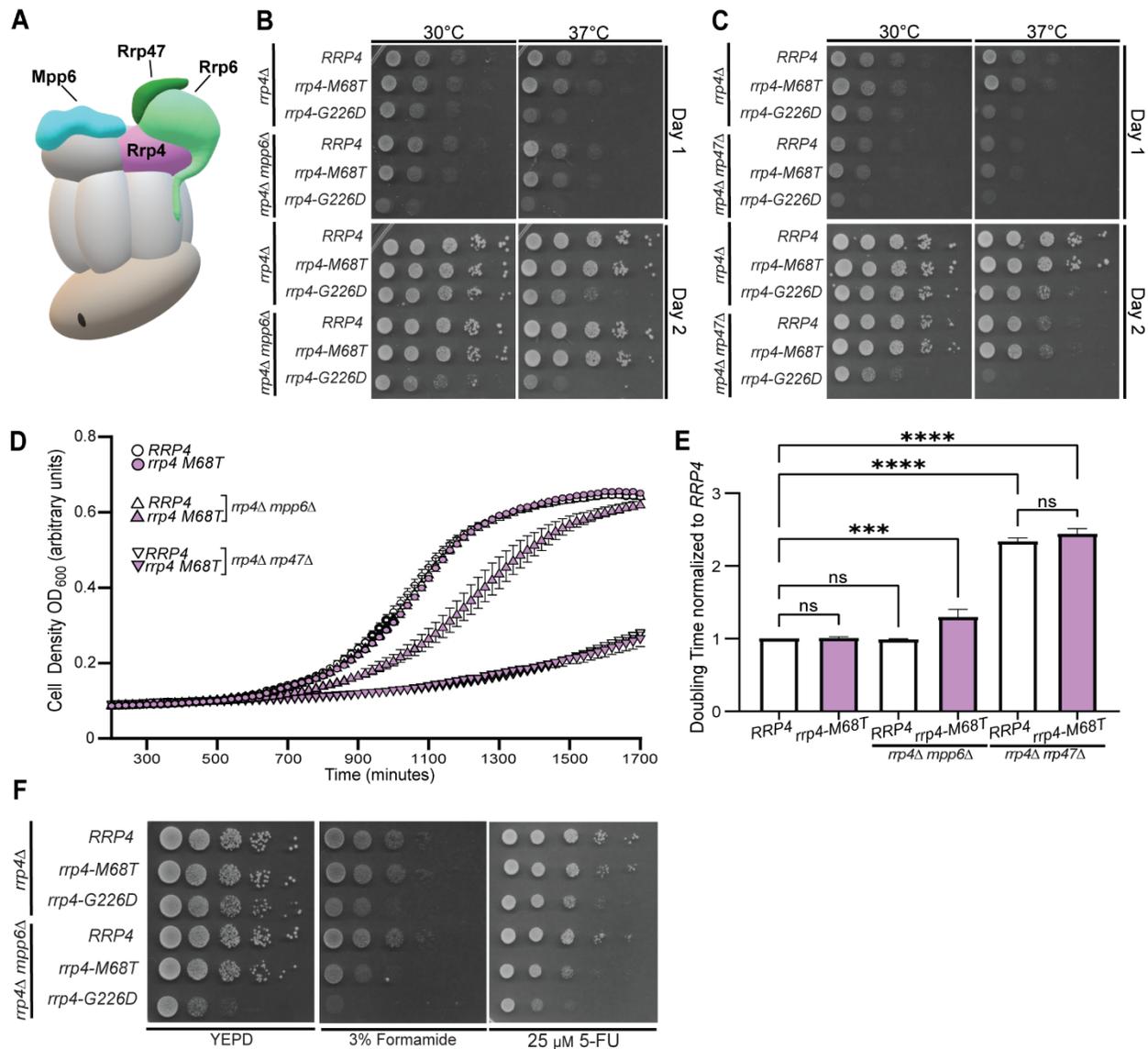


Figure 7. The *rrp4-M68T mpp6Δ* double mutant cells exhibits impaired growth that is exacerbated on drugs that impact RNA processing.

(A) Cartoon schematic of the budding yeast nuclear RNA exosome in complex with nuclear cofactors Mpp6 (turquoise) and Rrp6/47 (light green/dark green) [32]. Serial dilution growth assays of double mutant (B) *rrp4-M68T mpp6Δ* or (C) *rrp4-M68T rrp47Δ* cells at 30°C and 37°C. The double mutant cells (*rrp4Δ* with *mpp6Δ*, or *rrp47Δ*) containing control *RRP4* or *rrp4* variants

rrp4-M68T and *rrp4-G226D* plasmids were serially diluted, spotted onto selective solid media, and grown at the indicated temperatures for two days. The double mutant cells *rrp4-G226D mpp6Δ* and *rrp4-G226D rrp47Δ* were included as a comparative control and show growth defects as described previously [64]. Data shown are representative of three independent assays (n = 3). (D) and (E) Double mutant cells containing *rrp4-G226D* and *mpp6Δ* exhibit a statistically significant increase in doubling time in liquid culture. Double mutant cells (*rrp4Δ mpp6Δ* or *rrp4Δ rrp47Δ*) containing control *RRP4* or *rrp4-M68T* plasmids were diluted in selective media and grown at 37°C with optical density measurements used to assess cell density over time. Data shown is collected from four independent samples (n = 4). (E) Doubling time for each sample was quantified and normalized to the growth rate of control *RRP4* cells. All calculations were performed as described in *Materials and Methods*. Full liquid growth curves of both *rrp4-M68T mpp6Δ* and *rrp4-M68T rrp47Δ* mutant cells are shown in Supplemental Figure S5. (F) Double mutant cells *rrp4-M68T mpp6Δ* exhibit impaired growth on solid media containing drugs impacting RNA processing. The *rrp4Δ mpp6Δ* cells expressing *RRP4*, *rrp4-M68T* or *rrp4-G226D* were serially diluted, spotted onto solid YEPD media containing 3% formamide or selective media containing 25 μM fluorouracil (5-FU) and grown at 30°C for three days. The *rrp4-M68T mpp6Δ* cells show impaired growth when compared to *RRP4 mpp6Δ* cells. The *rrp4-G226D mpp6Δ* cells show exacerbated growth defects on 3% formamide and 25 μM 5-FU at 30°C. Data shown are representative of three independent assays (n = 3).

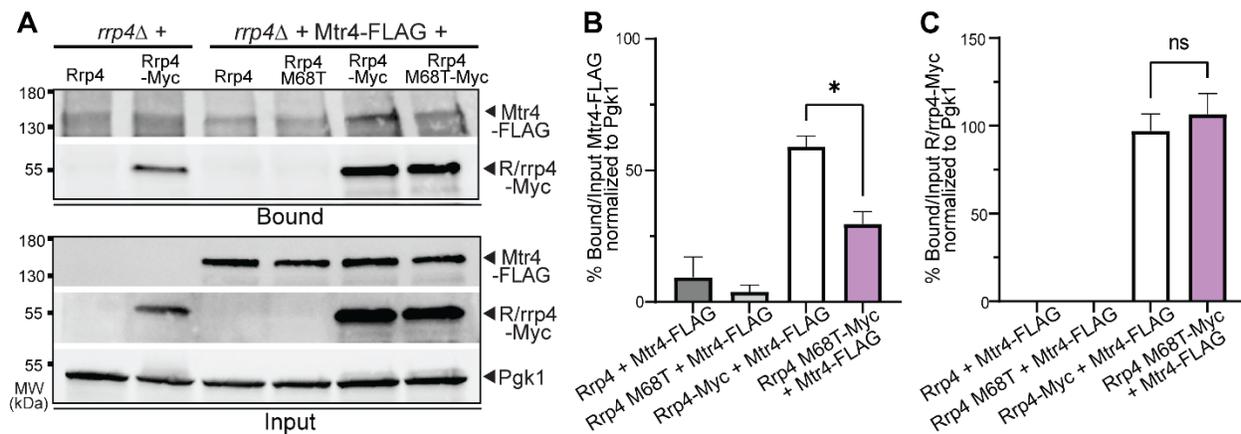
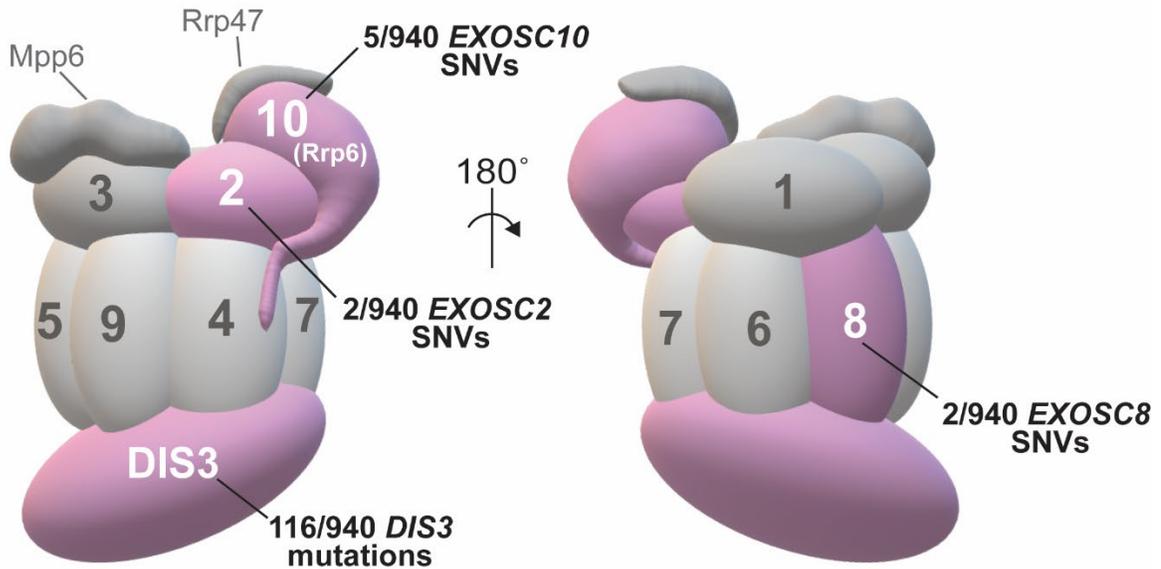


Figure 8. Rrp4 M68T shows decreased association with Mtr4 compared to wild-type Rrp4.

(A) The *rrp4* Δ cells co-expressing Rrp4-Myc variants and FLAG-tagged Mtr4 were grown at 30°C. Myc-tagged Rrp4 or Rrp4 M68T protein was immunoprecipitated from cleared lysate using anti-Myc beads and bound Mtr4-FLAG protein was detected by immunoblotting. The bound/input level of Mtr4-FLAG was detected with an anti-FLAG antibody and the bound/input level of Rrp4-Myc was detected with an anti-Myc antibody. The 3-phosphoglycerate kinase (Pgk1) serves as a loading control. (B) Quantitation of the percentage of bound to input Mtr4-FLAG co-immunoprecipitated with Rrp4-Myc or Rrp4 M68T-Myc normalized to Pgk1. The graph shows the mean percentage of bound Mtr4-FLAG normalized to unbound input. Error bars represent standard error of the mean. (C) Quantitation of percentage of bound to input Rrp4-Myc or Rrp4 M68T-Myc immunoprecipitated normalized to Pgk1. Error bars represent standard error of the mean. Statistical significance is denoted (* p -value ≤ 0.05 ; n.s. p -value ≥ 0.05). Data shown here collected were from two independent experiments ($n=2$). The co-immunoprecipitations were performed and quantitated as described in *Materials and Methods*.

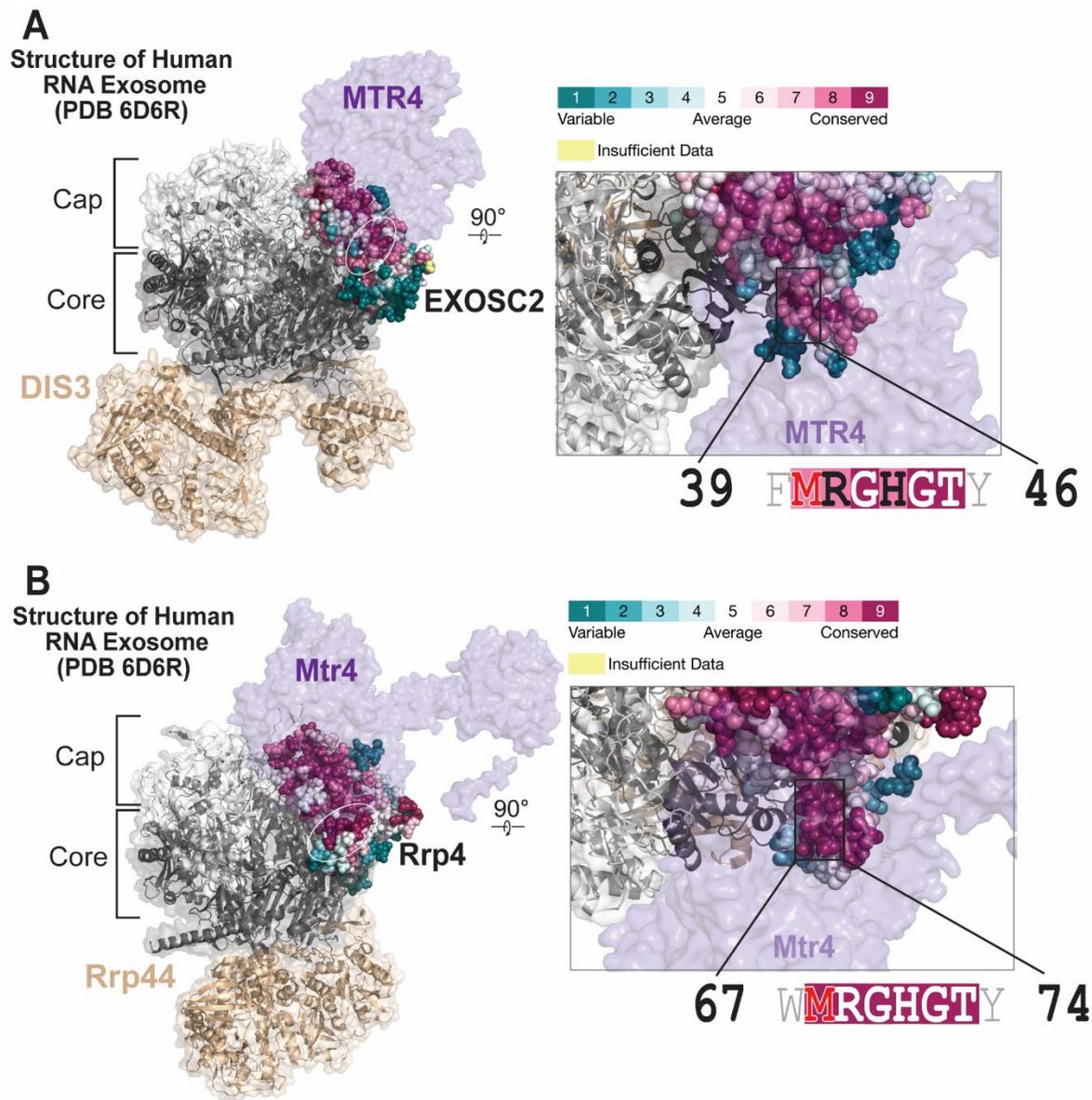
3.8 Chapter III Supplementary Materials



Supplementary Figure S1. A collection of mutations in RNA exosome subunit genes were identified in newly diagnosed multiple myeloma patients in the CoMMpass study.

Cartoon depicting the human nuclear RNA exosome ribonuclease complex. The nine structural subunits, EXOSC1-9, and the catalytic exo/endoribonuclease, DIS3, are labeled. The nuclear RNA exosome has an additional 3'-5' exonuclease, EXOSC10 (Rrp6 in yeast), that associates with the complex and aids in nuclear RNA targeting and processing. The associated nuclear cofactors Mpp6 and Rrp47 are also depicted. The ongoing longitudinal Multiple Myeloma Research Foundation (MMRF) study "Relating Clinical Outcomes in Multiple Myeloma to Personal Assessment of Genetic Profile" (CoMMpass) [ClinicalTrials.gov Identifier NCT01454297] identified mutations in *EXOSC2*, *EXOSC8*, *EXOSC10* and *DIS3* within multiple myeloma patients upon diagnosis (the encoded subunits are colored pink in the cartoon depiction). Whole genome sequencing was

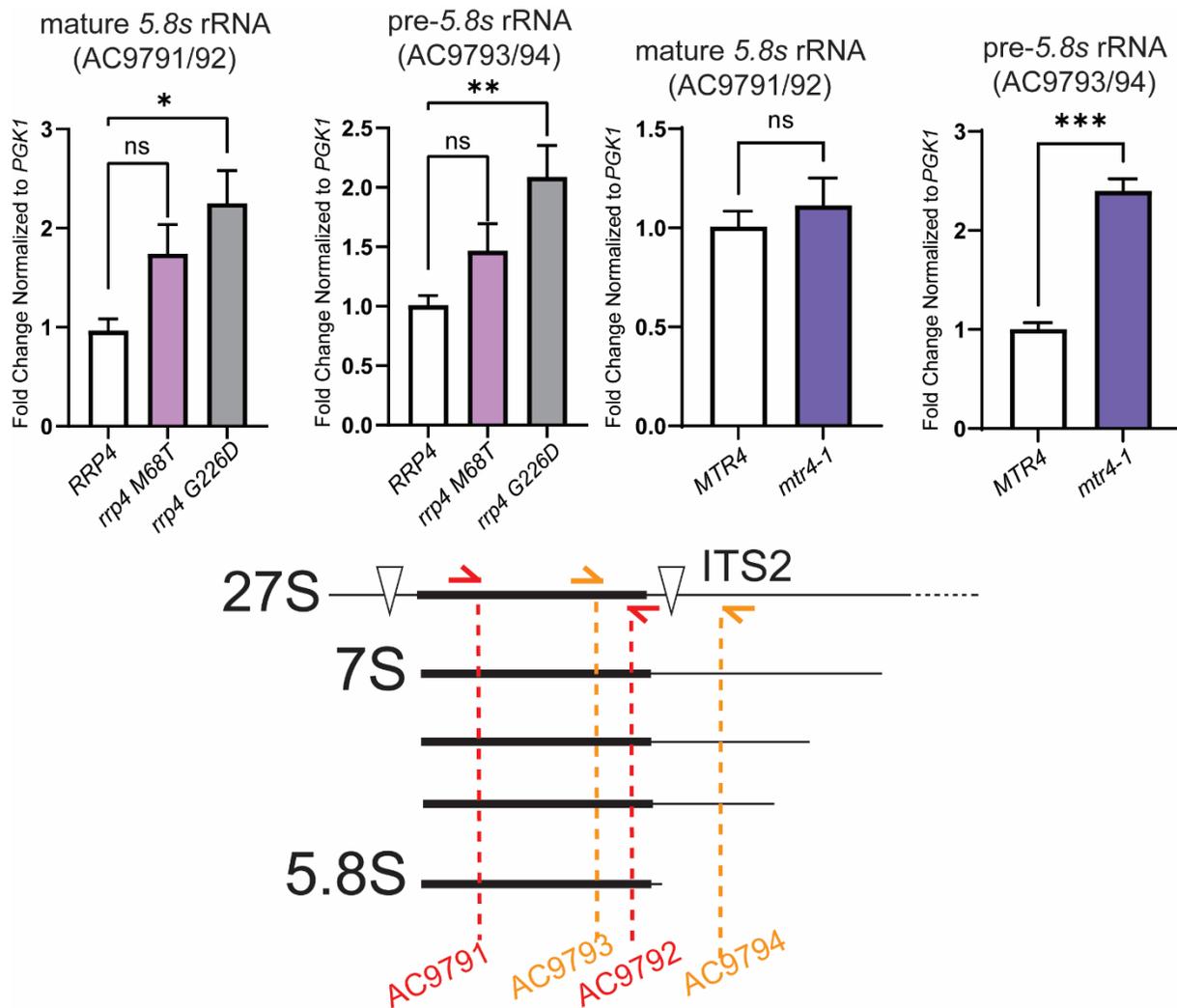
performed on 940 newly diagnosed patients through the CoMMpass study. Within that population, rare single nucleotide variants (SNVs) were identified in the structural RNA exosome cap and core genes, *EXOSC2* and *EXOSC8*. Additionally, SNVs were identified in *EXOSC10*. Mutations in *DIS3* were identified more commonly within newly diagnosed patients in CoMMpass. The number of newly diagnosed patients with SNVs or mutations in *EXOSC2*, *EXOSC8*, *EXOSC10* or *DIS3* identified through CoMMpass is listed in the figure next to the corresponding subunit.



Supplementary Figure S2. ConSurf analysis of EXOSC2 and Rrp4 reveals conservation at interface with MTR4/Mtr4.

The ConSurf server tool [69-71] was used to assess the conservation of both EXOSC2 (A) and Rrp4 (B). Residues are colored based on calculated conservation scores representing a relative measure of evolutionary conservation. Conservation scores range from 1 (blue) to 9 (pink).

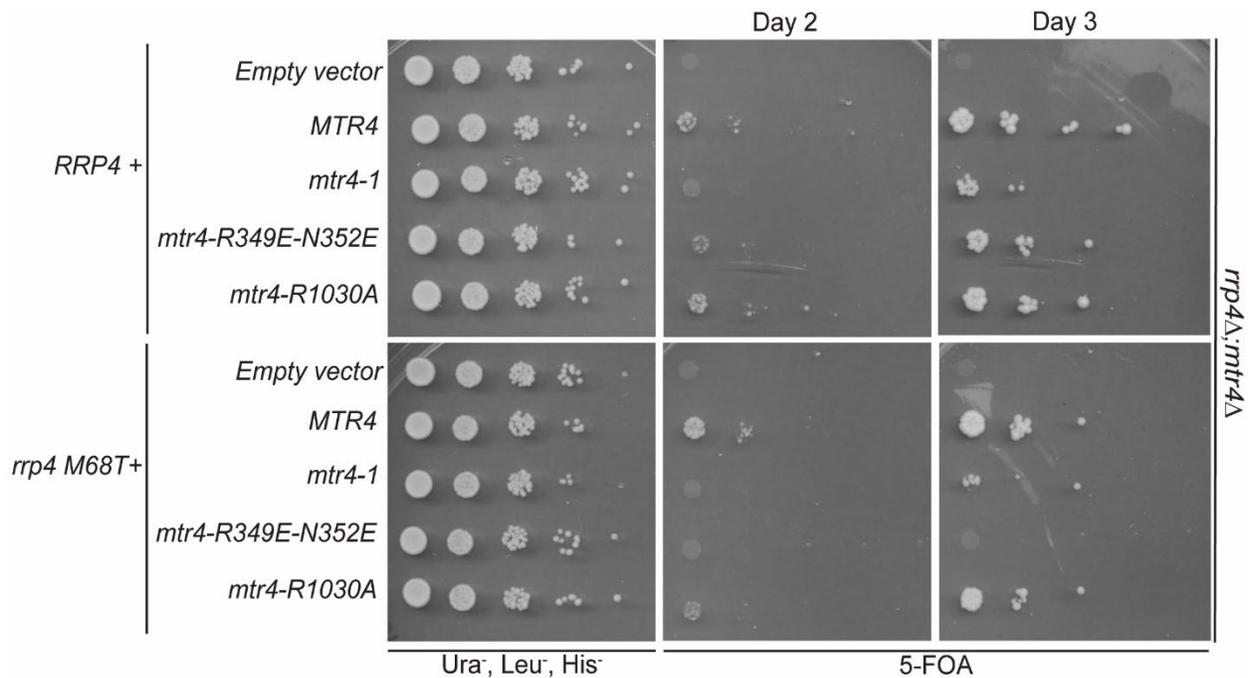
Calculated conservation scores that do not pass statistical tests are marked as having insufficient data and are colored yellow in the structure. Conservation score colors are mapped onto structures of EXOSC2 in the mammalian RNA exosome (PDB 6D6R) and Rrp4 in the budding yeast RNA exosome (PDB 6FSZ). Both structures include the RNA helicase MTR4/Mtr4 (purple). Zoomed insets show the conservation of the region of EXOSC2 that includes Met40 and the corresponding Rrp4 Met68.



Supplementary Figure S3. The steady-state level of mature and precursor 5.8S rRNA in *rrp4-M68T* cells is similar to wild-type, control cells.

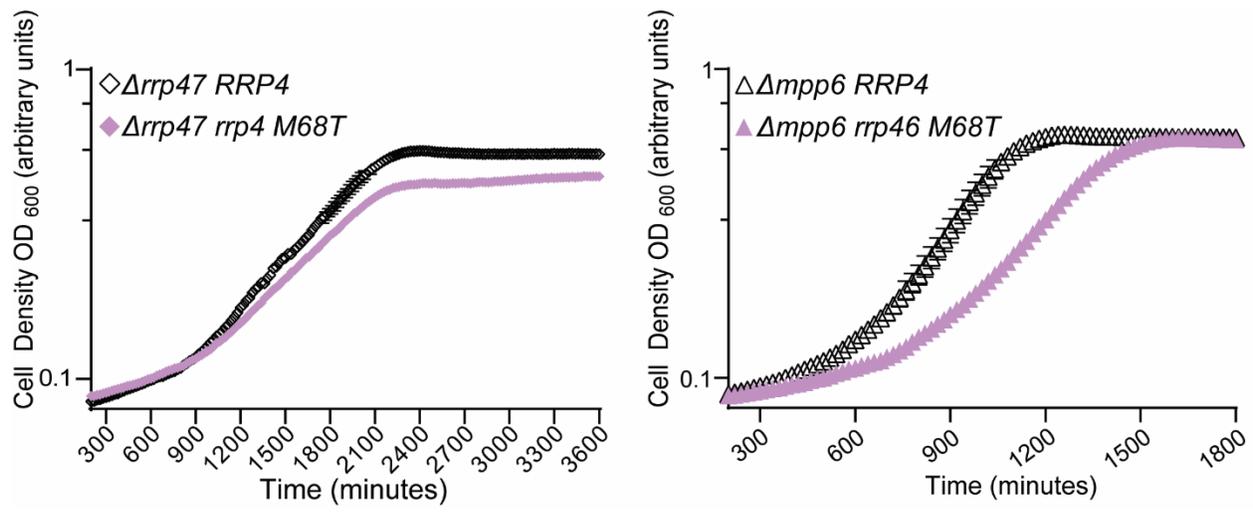
Total RNA from mutant cells (*rrp4-M68T*, *rrp4-G226D*, *mtr4-1*) and the corresponding wild-type control cells (*RRP4* and *MTR4*) grown at 37°C was extracted analyzed by RT-qPCR with primers that amplifies the mature 5.8S rRNA (AC9791/9792; red) or with primers that flank the 5.8S-ITS2 junction (AC9793/9794; orange) to detect 7S pre-rRNA. The 7S pre-rRNA is normally processed to mature 5.8S rRNA by 3'-5' decay of the internal transcribed spacer 2 (ITS2) via the nuclear

RNA exosome [3, 9]. The simplified schematics to the right illustrate the processing steps of 7S rRNA precursor following endonucleolytic cleavage from the larger 27S precursor (indicated by white triangles). The locations of the primer sets are denoted on the simplified schematic. The steady-state level of mature 5.8S and pre-5.8S rRNA is not significantly increased in *rrp4-M68T* cells (denoted in pink) compared to control *RRP4* cells. The steady-state levels of these rRNA species are significantly increased in *rrp4-G226D* cells (denoted in gray) as previously reported [64]. Furthermore, the steady-state level of pre-5.8S rRNA is significantly increased in the *mtr4-1* cells when compared to control *MTR4* cells while the level of mature 5.8S rRNA is not significantly different between the *mtr4-1* and *MTR4* cells (denoted in purple). RNA isolation and RT-qPCR were performed as described in *Materials and Methods*. Statistical significance of the RNA levels in *rrp4* variant cells relative to *RRP4* cells and in the *mtr4-1* cells relative to *MTR4* cells is denoted by an asterisk (* p -value ≤ 0.05 ; ** p -value ≤ 0.01 , *** p -value ≤ 0.001).



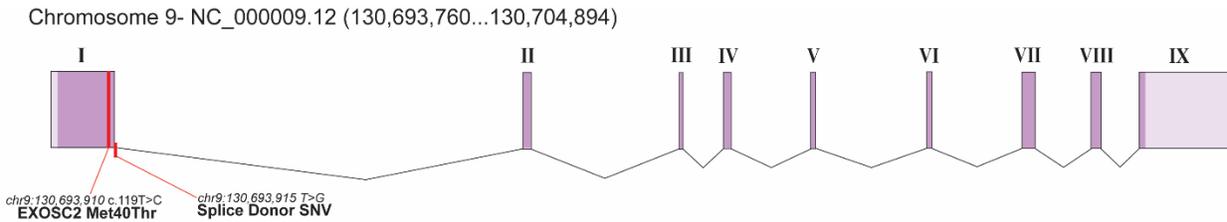
Supplementary Figure S4. Synthetical lethality of either *rrp4-M68T mtr4-R349E-N352E* and *rrp4-M68T mtr4-R1030A* double mutant cells is rescued by wild-type plasmid.

The *rrp4Δ mtr4Δ* cells containing *RRP4 URA3* and *MTR4 URA3* maintenance plasmids were transformed with empty vectors, or the *MTR4/mtr4 HIS3* plasmids and the *RRP4/rrp4 LEU* plasmids. Cells were grown overnight and serially diluted and spotted onto Ura⁻ Leu⁻ His⁻ minimal media plates, which select for cells that contain *URA3* maintenance plasmids, the *RRP4/rrp4 LEU2* plasmid, and the *MTR4/mtr4 HIS3* plasmid. Cells were also spotted onto 5-FOA Leu⁻ minimal media plates, which selects for cells that lack the *URA3* maintenance plasmids and contain only the *RRP4/rrp4 LEU2* and *MTR4/mtr4 HIS3* plasmids. The plates were incubated at 30°C for 3 days.



Supplementary Figure S5. Extended liquid growth curve of *rrp4-M68T mpp6Δ* and *rrp4-M68T rrp47Δ* cells.

The *rrp4Δ mpp6Δ* or *rrp4Δ rrp47Δ* cells expressing *RRP4* or *rrp4-M68T* were grown in liquid media at 37°C with optical density measurement used to assess cell density over time. Data shown are collected from four independent samples (n = 4). These growth curves are the source of the data displayed in Figure 6D carried out for a longer time course and used to quantify the doubling time presented in Figure 6E.



Supplementary Figure S6. Chr9- NC_000009.12 (130,693,760...130,704,894) schematic with multiple myeloma patient *EXOSC2* mutations.

Gene schematic depicting the *EXOSC2* locus. The gene contains nine exons labeled in roman numerals. Coding sequence and the 5' and 3' UTRs are color coded in pink and light pink, respectively. The multiple myeloma patient mutation location is depicted in red. The missense mutation *chr9:130,693,910 c.119 T>C* is in exon I and results in the *EXOSC2* Met40Thr substitution. The single nucleotide variant (SNV) *chr9:130,693,915 T>G* is a splice donor mutation within the first intron. This splice donor SNV is reported in NIH dsSNP (rs1430887213).

Table S1. *S. cerevisiae* Strains and Plasmids used in this study.

Strain/Plasmid	Description	Source
<i>rrp4Δ</i> (yAV1103)	<i>MATa; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; RRP4::neoMX (G418+); [RRP4; URA3]</i>	Losh 2018[74]
<i>rrp4Δmpp6Δ</i> (ACY2471)	<i>MATa; ura3Δ0; leu2Δ0; his3Δ1; RRP4::neoMX (G418+); [RRP4; URA3]; MPP6::natMX4</i>	Sterrett Enyenihi et al. 2021
<i>rrp4Δrrp47Δ</i> (ACY2474)	<i>MATa; ura3Δ0; leu2Δ0; his3Δ1; RRP4::neoMX (G418+); [RRP4; URA3]; LRP1::natMX4</i>	Sterrett Enyenihi et al. 2021
<i>RRP45-TAP</i> (ACY2789)	<i>MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; RRP45-TAP:HIS3MX6</i>	Ghaemmaghami et al. 2009
<i>rrp4Δmtr4Δ</i> (ACY2536)	<i>MATa; ura3-; leu2-; his3-; trp1-; LYS+; GAL+; ADE+; MTR4::natMX4; [pAC3714; MTR4; RRP4; URA3; CEN]; RRP4::neoMX</i>	Sterrett Enyenihi et al. 2021
<i>mtr4Δ</i> (ACY2532)	<i>MATa; ura3-; leu2-; his3-; trp1-; LYS+; GAL+; ADE+; MTR4::natMX4; [pAC3714; MTR4; RRP4; URA3; CEN]</i>	This Study
pRS315 (pAC3)	<i>CEN6, LEU2, ampR</i>	Sikorski and Hieter 1989
pRS313 (pAC1)	<i>CEN6, HIS3, amp^R</i>	Sikorski and Hieter 1989
pAC3656	<i>RRP4-Native 3' UTR in pRS315, CEN6, LEU2, ampR</i>	Sterrett Enyenihi et al. 2021
pAC3669	<i>RRP4-2xMyc-Native 3' UTR in pRS315, CEN6, LEU2, ampR</i>	Sterrett Enyenihi et al. 2021
pAC3659	<i>rrp4-G226D-Native 3' UTR in pRS315, CEN6, LEU2, ampR</i>	Sterrett Enyenihi et al. 2021
pAC3672	<i>rrp4-G226D-2xMyc-Native 3' UTR in pRS315, CEN6, LEU2, ampR</i>	Sterrett Enyenihi et al. 2021
pAC3714	<i>MTR4, RRP4, CEN6, URA3, amp^R</i>	Sterrett Enyenihi et al. 2021
pAC3719	<i>MTR4-2xFLAG, CEN6, HIS, amp^R</i>	Sterrett Enyenihi et al. 2021
pAC4096	<i>MTR4-Native 3'UTR in pRS313, CEN6, HIS3, ampR</i>	Sterrett Enyenihi et al. 2021
pAC4099	<i>mtr4-F7A-F10A-Native 3'UTR in pRS313, CEN6, HIS3, ampR</i>	Sterrett Enyenihi et al. 2021

pAC4103	<i>mtr4-1-(mtr4-C942Y)-Native 3'UTR</i> in pRS313, <i>CEN6, HIS3, ampR</i>	This Study
pAC4104	<i>mtr4-R1030A-Native 3'UTR</i> in pRS313, <i>CEN6, HIS3, ampR</i>	This Study
pAC4105	<i>mtr4-E1033W-Native 3'UTR</i> in pRS313, <i>CEN6, HIS3, ampR</i>	This Study
pAC4206	<i>rrp4-M68T-Native 3' UTR</i> in pRS315, <i>CEN6, LEU2, ampR</i>	This Study
pAC4207	<i>rrp4-M68T-2xMyc-Native 3' UTR</i> in pRS315, <i>CEN6, LEU2, ampR</i>	This Study

Table S2. DNA Oligonucleotides employed for RT-qPCR

Description	Sequence (5'-3')	Name
<i>INO1 mRNA</i> Fwd	TTGGACTGCAAATACTGAGAGG	AC9303
<i>INO1 mRNA</i> Rev	AAGATCGTGGAAGGAGCAATC	AC9302
<i>CUT501 ncRNA</i> Fwd	GCTAGCACCTGTTGCTGTAAT	AC9255
<i>CUT501 ncRNA</i> Rev	GGTTCAACGTTGCAGGATCT	AC9254
<i>CUT770 ncRNA</i> Fwd	AAACAACCCGCTAGTGTGAC	AC9262
<i>CUT770 ncRNA</i> Rev	AGAGCAACTCACTGCAAAGG	AC9263
<i>CUT896 ncRNA</i> Fwd	CCCAGAGGCAAAGATGTTAAGT	AC9257
<i>CUT896 ncRNA</i> Rev	ATCAGCAGGTGTCATGTTACAG	AC9256
<i>pre-TLC1 ncRNA</i> Fwd	CCGCCTATCCTCGTCATGAAC	AC7594
<i>pre-TLC1 ncRNA</i> Rev	GTATTGTAGAAATCGCGCGTAC	AC7593
mature <i>TLC1 ncRNA</i> Fwd	AAGGCAAGGGTGTCTTTCT	AC6420
mature <i>TLC1 ncRNA</i> Rev	TTCCGCTTGAAAATAATGC	AC6421
3' extended <i>U4 snRNA</i> Fwd	ATCCTTATGCACGGGAAATACG	AC5722
3' extended <i>U4 snRNA</i> Rev	AAAGAATGAATATCGGTAATG	AC5723
3' extended <i>snR33 snoRNA</i> Fwd	AAGCGACCTTTCTTCGCA	AC9787
3' extended <i>snR33 snoRNA</i> Rev	TTCGCTTCTGGTTACTGCAA	AC9788
<i>5.8S rRNA</i> mature Fwd	CAACAACGGATCTCTTGTTCT	AC9791
<i>5.8S rRNA</i> mature Rev	GAAATGACGCTCAAACAGGCA	AC9792
<i>5.8S-ITS2 rRNA marginal</i> Fwd	CGAATCTTTGAACGCACATTGC	AC9793
<i>5.8S rRNA precursor 3'</i> Rev	GGAAATGACGCTCAAACAGG	AC9794
<i>ALG9 mRNA</i> Fwd	CACGGATAGTGGCTTTGGTGAACAATTAC	AC5067
<i>ALG9 mRNA</i> Rev	TATGATTATCTGGCAGCAGGAAAGAACTTGGG	AC5068
<i>PGK1 mRNA</i> Fwd	CTGCTTTGCCAACCATCAAGT	AC2307
<i>PGK1 mRNA</i> Rev	GCAACTGGAGCCAAAGAGTATTTT	AC2308

3.9 References

1. Zinder, J.C. and C.D. Lima, *Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors*. Genes & Development, 2017. **31**(2): p. 88-100.
2. Schneider, C. and D. Tollervey, *Threading the barrel of the RNA exosome*. Trends in Biochemical Sciences, 2013. **38**(10): p. 485-493.
3. Mitchell, P., E. Petfalski, and D. Tollervey, *The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism*. Genes Dev, 1996. **10**(4): p. 502-13.
4. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'->5' exoribonucleases*. Cell, 1997. **91**(4): p. 457-466.
5. Lorentzen, E., et al., *RNA channelling by the archaeal exosome*. Embo Reports, 2007. **8**(5): p. 470-476.
6. Hou, D., M. Ruiz, and E.D. Andrulis, *The ribonuclease Dis3 is an essential regulator of the developmental transcriptome*. BMC Genomics, 2012. **13**.
7. Lim, S.J., et al., *Genome-wide localization of exosome components to active promoters and chromatin insulators in Drosophila*. Nucleic Acids Res, 2013. **41**(5): p. 2963-80.
8. Pefanis, E., et al., *Noncoding RNA transcription targets AID to divergently transcribed loci in B cells*. Nature, 2014. **514**(7522): p. 389-93.
9. Allmang, C., et al., *Functions of the exosome in rRNA, snoRNA and snRNA synthesis*. Embo Journal, 1999. **18**(19): p. 5399-5410.
10. van Hoof, A., P. Lennertz, and R. Parker, *Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs*. Molecular and cellular biology, 2000. **20**(2): p. 441-452.
11. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 227-239.
12. Fasken, M.B., et al., *The RNA Exosome and Human Disease*. Methods Mol Biol, 2020. **2062**: p. 3-33.
13. Wyers, F., et al., *Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase*. Cell, 2005. **121**(5): p. 725-737.
14. Parker, R., *RNA Degradation in Saccharomyces cerevisiae*. Genetics, 2012. **191**(3): p. 671-702.

15. Schneider, C., et al., *Transcriptome-wide Analysis of Exosome Targets*. Molecular Cell, 2012. **48**(3): p. 422-433.
16. Moore, M.J. and N.J. Proudfoot, *Pre-mRNA processing reaches back to transcription and ahead to translation*. Cell, 2009. **136**(4): p. 688-700.
17. Preker, P., et al., *RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters*. Science, 2008. **322**(5909): p. 1851-1854.
18. Briggs, M.W., K.T. Burkard, and J.S. Butler, *Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation*. J Biol Chem, 1998. **273**(21): p. 13255-63.
19. Wasmuth, E.V., K. Januszyk, and C.D. Lima, *Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA*. Nature, 2014. **511**(7510): p. 435-9.
20. Makino, D.L., M. Baumgaertner, and E. Conti, *Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex*. Nature, 2013. **495**(7439): p. 70-75.
21. Schuller Jan, M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. Science, 2018. **360**(6385): p. 219-222.
22. Weick, E.M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. Cell, 2018. **173**(7): p. 1663-1677 e21.
23. Ogami, K., Y. Chen, and J.L. Manley, *RNA surveillance by the nuclear RNA exosome: mechanisms and significance*. Noncoding RNA, 2018. **4**(1).
24. de la Cruz, J., et al., *Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae*. Embo j, 1998. **17**(4): p. 1128-40.
25. LaCava, J., et al., *RNA degradation by the exosome is promoted by a nuclear polyadenylation complex*. Cell, 2005. **121**(5): p. 713-724.
26. Vaňáčková, Š., et al., *A New Yeast Poly(A) Polymerase Complex Involved in RNA Quality Control*. PLOS Biology, 2005. **3**(6): p. e189.
27. Mitchell, P., et al., *Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs*. Mol Cell Biol, 2003. **23**(19): p. 6982-92.
28. Milligan, L., et al., *A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts*. Mol Cell Biol, 2008. **28**(17): p. 5446-57.
29. Schuch, B., et al., *The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase*. Embo Journal, 2014. **33**(23): p. 2829-2846.

30. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. Cell Reports, 2017. **20**(10): p. 2279-2286.
31. Wasmuth, E.V., et al., *Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase*. Elife, 2017. **6**: p. 24.
32. Schuller, J.M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. Science, 2018. **360**(6385): p. 219-222.
33. Taylor, L.L., et al., *The Mtr4 ratchet helix and arch domain both function to promote RNA unwinding*. Nucleic acids research, 2014. **42**(22): p. 13861-13872.
34. Weir, J.R., et al., *Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance*. Proc Natl Acad Sci U S A, 2010. **107**(27): p. 12139-44.
35. Stuparevic, I., et al., *Cotranscriptional Recruitment of RNA Exosome Cofactors Rrp47p and Mpp6p and Two Distinct Trf-Air-Mtr4 Polyadenylation (TRAMP) Complexes Assists the Exonuclease Rrp6p in the Targeting and Degradation of an Aberrant Messenger Ribonucleoprotein Particle (mRNP) in Yeast*. Journal of Biological Chemistry, 2013. **288**(44): p. 31816-31829.
36. Falk, S., et al., *The molecular architecture of the TRAMP complex reveals the organization and interplay of its two catalytic activities*. Mol Cell, 2014. **55**(6): p. 856-867.
37. Houseley, J. and D. Tollervey, *The many pathways of RNA degradation*. Cell, 2009. **136**(4): p. 763-76.
38. Houseley, J. and D. Tollervey, *The nuclear RNA surveillance machinery: the link between ncRNAs and genome structure in budding yeast?* Biochim Biophys Acta, 2008. **1779**(4): p. 239-46.
39. Lubas, M., et al., *Interaction profiling identifies the human nuclear exosome targeting complex*. Mol Cell, 2011. **43**(4): p. 624-37.
40. Liang, S., et al., *A DEAD-box-family protein is required for nucleocytoplasmic transport of yeast mRNA*. Mol Cell Biol, 1996. **16**(9): p. 5139-46.
41. Kadowaki, T., et al., *Mutations in nucleolar proteins lead to nucleolar accumulation of polyA+ RNA in Saccharomyces cerevisiae*. Mol Biol Cell, 1995. **6**(9): p. 1103-10.
42. Kadowaki, T., et al., *Isolation and characterization of Saccharomyces cerevisiae mRNA transport-defective (mtr) mutants*. J Cell Biol, 1994. **126**(3): p. 649-59.
43. Cervelli, T. and A. Galli, *Yeast as a Tool to Understand the Significance of Human Disease-Associated Gene Variants*. Genes (Basel), 2021. **12**(9).

44. Staals, R.H., et al., *Dis3-like 1: a novel exoribonuclease associated with the human exosome*. *Embo j*, 2010. **29**(14): p. 2358-67.
45. Lohr, J.G., et al., *Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy*. *Cancer Cell*, 2014. **25**(1): p. 91-101.
46. Chapman, M.A., et al., *Initial genome sequencing and analysis of multiple myeloma*. *Nature*, 2011. **471**(7339): p. 467-72.
47. Alexander, D.D., et al., *Multiple myeloma: a review of the epidemiologic literature*. *Int J Cancer*, 2007. **120 Suppl 12**: p. 40-61.
48. Laubach, J., P. Richardson, and K. Anderson, *Multiple myeloma*. *Annu Rev Med*, 2011. **62**: p. 249-64.
49. Weissbach, S., et al., *The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma*. *Br J Haematol*, 2015. **169**(1): p. 57-70.
50. Tomecki, R., et al., *Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target*. *Nucleic Acids Res*, 2014. **42**(2): p. 1270-90.
51. Boyle, E.M., et al., *BRAF and DIS3 Mutations Associate with Adverse Outcome in a Long-term Follow-up of Patients with Multiple Myeloma*. *Clin Cancer Res*, 2020. **26**(10): p. 2422-2432.
52. Eggens, V.R.C., et al., *EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations*. *Orphanet Journal of Rare Diseases*, 2014. **9**(1): p. 23.
53. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. *Hum Mol Genet*, 2020. **29**(13): p. 2218-2239.
54. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. *Nature Communications*, 2014. **5**.
55. Di Donato, N., et al., *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*. *Journal of Medical Genetics*, 2016. **53**(6): p. 419-425.
56. Burns, D., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. *Neuromuscul Disord*, 2017. **27**: p. S38.

57. Somashekar, P.H., et al., *Bi-allelic missense variant, p.Ser35Leu in EXOSC1 is associated with pontocerebellar hypoplasia*. Clin Genet, 2021. **99**(4): p. 594-600.
58. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nat Genet, 2012. **44**(6): p. 704-8.
59. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
60. Yang, X., et al., *Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis*. Human Molecular Genetics, 2019. **29**(4): p. 541-553.
61. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. Genetics, 2017. **205**(1): p. 221-+.
62. Gillespie, A., et al., *Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in S. cerevisiae*. RNA, 2017. **23**(4): p. 466-472.
63. de Amorim, J., Slavotinek, A., Fasken, M.B., Corbett, A.H., Morton, D.J., *Modeling Pathogenic Variants in the RNA Exosome*. RNA & Disease, 2020. **7**.
64. Sterrett, M.C., et al., *A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex*. RNA, 2021. **27**(9): p. 1046-1067.
65. Morton, D.J., et al., *A Drosophila model of Pontocerebellar Hypoplasia reveals a critical role for the RNA exosome in neurons*. PLoS Genet, 2020. **16**(7): p. e1008901.
66. Barwick, B.G., et al., *Multiple myeloma immunoglobulin lambda translocations portend poor prognosis*. Nature Communications, 2019. **10**(1): p. 1911.
67. Adams, A., et al., *Methods in Yeast Genetics*. 1997, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
68. Rodrigues, C.H.M., et al., *mCSM-PPI2: predicting the effects of mutations on protein-protein interactions*. Nucleic Acids Res, 2019. **47**(W1): p. W338-w344.
69. Celniker, G., et al., *ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function*. Israel Journal of Chemistry, 2013. **53**(3-4): p. 199-206.
70. Ashkenazy, H., et al., *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules*. Nucleic Acids Research, 2016. **44**(W1): p. W344-W350.

71. Ashkenazy, H., et al., *ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids*. Nucleic Acids Research, 2010. **38**(suppl_2): p. W529-W533.
72. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*. Second ed. 1989, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
73. Schaeffer, D., et al., *The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities*. Nat Struct Mol Biol, 2009. **16**(1): p. 56-62.
74. Losh, J., *Identifying Subunit Organization and Function of the Nuclear RNA Exosome Machinery*. 2018, University of Texas: UT GSBS Dissertations and Theses.
75. Ghaemmaghami, S., et al., *Global analysis of protein expression in yeast*. Nature, 2003. **425**(6959): p. 737-41.
76. Da, B., D. Dawson, and T. Stearns, *Methods In Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. 2000.
77. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
78. Bergsagel, P.L., et al., *IgH translocations in multiple myeloma: a nearly universal event that rarely involves c-myc*. Curr Top Microbiol Immunol, 1997. **224**: p. 283-7.
79. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae*. Genetics, 1989. **122**(1): p. 19-27.
80. Hoskins, J. and J. Scott Butler, *Evidence for distinct DNA- and RNA-based mechanisms of 5-fluorouracil cytotoxicity in Saccharomyces cerevisiae*. Yeast, 2007. **24**(10): p. 861-70.
81. Giaever, G., et al., *Chemogenomic profiling: identifying the functional interactions of small molecules in yeast*. Proc Natl Acad Sci U S A, 2004. **101**(3): p. 793-8.
82. Hoyos-Manchado, R., et al., *RNA metabolism is the primary target of formamide in vivo*. Scientific Reports, 2017. **7**(1): p. 15895.
83. Slater, M.L., *Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle*. J Bacteriol, 1973. **113**(1): p. 263-70.
84. Suzuki, H., et al., *On the mechanism of action of bleomycin. Strand scission of DNA caused by bleomycin and its binding to DNA in vitro*. J Antibiot (Tokyo), 1970. **23**(10): p. 473-80.

85. Coy, S., et al., *The Sm complex is required for the processing of non-coding RNAs by the exosome*. PloS one, 2013. **8**(6): p. e65606-e65606.
86. Houseley, J., J. LaCava, and D. Tollervey, *RNA-quality control by the exosome*. Nat Rev Mol Cell Biol, 2006. **7**(7): p. 529-39.
87. Donahue, T.F. and S.A. Henry, *myo-Inositol-1-phosphate synthase. Characteristics of the enzyme and identification of its structural gene in yeast*. J Biol Chem, 1981. **256**(13): p. 7077-85.
88. Klig, L.S. and S.A. Henry, *Isolation of the yeast INO1 gene: located on an autonomously replicating plasmid, the gene is fully regulated*. Proc Natl Acad Sci U S A, 1984. **81**(12): p. 3816-20.
89. Delan-Forino, C., C. Schneider, and D. Tollervey, *Transcriptome-wide analysis of alternative routes for RNA substrates into the exosome complex*. PLOS Genetics, 2017. **13**(3): p. e1006699.
90. Rodríguez-Galán, O., et al., *Immature large ribosomal subunits containing the 7S pre-rRNA can engage in translation in Saccharomyces cerevisiae*. RNA biology, 2015. **12**(8): p. 838-846.
91. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nat Rev Mol Cell Biol, 2016. **17**(4): p. 227-39.
92. Belair, C., et al., *The RNA exosome nuclease complex regulates human embryonic stem cell differentiation*. J Cell Biol, 2019.
93. Belair, C., S. Sim, and S.L. Wolin, *Noncoding RNA Surveillance: The Ends Justify the Means*. Chem Rev, 2018. **118**(8): p. 4422-4447.
94. Anderson, J.T. and X. Wang, *Nuclear RNA surveillance: no sign of substrates tailing off*. Crit Rev Biochem Mol Biol, 2009. **44**(1): p. 16-24.
95. Lim, J., et al., *Nuclear Proximity of Mtr4 to RNA Exosome Restricts DNA Mutational Asymmetry*. Cell, 2017. **169**(3): p. 523-537 e15.
96. Oddone, A., et al., *Structural and biochemical characterization of the yeast exosome component Rrp40*. EMBO reports, 2007. **8**(1): p. 63-69.
97. Büttner, K., S. Nehring, and K.P. Hopfner, *Structural basis for DNA duplex separation by a superfamily-2 helicase*. Nat Struct Mol Biol, 2007. **14**(7): p. 647-52.
98. Jackson, R.N., et al., *The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing*. EMBO J, 2010. **29**(13): p. 2205-16.
99. Cherry, J.M., et al., *Saccharomyces Genome Database: the genomics resource of budding yeast*. Nucleic Acids Res, 2012. **40**(Database issue): p. D700-5.

Chapter IV. Comparative analysis of disease-linked amino acid substitutions in the RNA exosome modeled in *S. cerevisiae* reveal functional consequences in translation.

The research and data presented in this chapter will be included in a future manuscript entitled “Comparative analysis of disease-linked amino acid substitutions in the RNA exosome modeled in *S. cerevisiae* reveal functional consequences in translation.”

Sterrett, M. C., M. B. Fasken, S. Khoshnevis, A. H. Corbett and H. Ghalei (2023). “Comparative analysis of disease-linked amino acid substitutions in the RNA exosome modeled in *S. cerevisiae* reveal functional consequences in translation.” *In preparation*.

Author Contributions:

M.C Sterrett generated all the data and performed all analyses. M.B. Fasken assisted in the original construction of reagents used in Chapter IV.

4.1 Abstract

RNA exosomopathies is a growing family of diseases that are linked to missense mutations in genes encoding structural subunits of the RNA exosome. The RNA exosome is an evolutionarily conserved, 10-subunit exoribonuclease complex that consists of a 3-subunit cap, a 6-subunit, barrel-shaped, core and a catalytic base subunit. RNA exosomopathy disease pathologies are diverse, ranging from neurological defects to developmental disorders. The diversity of the RNA exosomopathy pathologies suggest that the different missense mutations in structural genes result in distinct *in vivo* consequences. To investigate these functional consequences and distinguish whether they are unique to each RNA exosomopathy mutation, we generated a collection of *in vivo* models using budding yeast by introducing pathogenic missense mutations in orthologous *S. cerevisiae* genes. We then performed a comparative RNA-Seq analysis to assess broad transcriptomic changes resulting in each mutant model. From these data we observe three of our mutant models—*rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H*, which model mutations in the structural genes *EXOSC2*, *EXOSC3* and *EXOSC5*—have the largest transcriptomic differences. Further analyses of these changes in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* reveal shared and distinct transcripts that are affected in the three mutant models. Gene Ontology (GO) analyses suggest similar impacts on biological processes across the three *rrp* mutant models, with shared decreased transcripts enriched in metabolic/biosynthetic processes and shared increased transcripts enriched in translation or ribosomal RNA modification/processing pathways. These results provide the first comparative RNA-Seq analysis of this number of RNA exosomopathy mutant models. Furthermore, these data suggest that different RNA exosomopathy mutations result in *in vivo* consequences that are both unique and shared across each variant, providing more insight into the biology that may underlie each distinct pathology.

4.2 Introduction

Though all cells within the human body contain the same genetic information, differential expression allows for cells to have varied functions, giving way to complex organization and tissues. Key to this cell-specific genetic expression are post-transcriptional events that regulate RNA levels. Cellular levels of both coding and non-coding RNAs are achieved through a delicate balance of transcription and decay. This balance is fine-tuned through post-transcriptional events that include precise processing, regulated decay and quality control surveillance [1]. These post-transcriptional events are critical to define the proteome in both time and space.

The RNA exosome is an abundant, essential cellular machine that is a critical mediator of both RNA processing and decay. This molecular machine is a multi-subunit complex composed of ten structural subunits and a catalytic 3'-5' exo-endoribonuclease (DIS3 [human]; Rrp44 [yeast]) [2, 3]. The subunits of the RNA exosome are highly conserved with many first identified in *Saccharomyces cerevisiae* in a screen for ribosomal RNA processing (*rrp*) mutants [3, 4]. As illustrated in Figure 1A, the 9-subunit structural core of the RNA exosome is composed of three S1/KH cap subunits (EXOSC1/2/3[human]; Csl4/Rrp4/Rrp40[yeast]) and a lower ring of six PH-like subunits (EXOSC4/7/8/9/5/6[human]; Rrp41/Rrp42/Rrp43/Rrp45/Rrp46/Mtr3[yeast]). This 9-subunit core forms a channel through which RNA substrates can be threaded to the associated catalytic DIS3 for processing or degradation (Figure 1A). Structural studies of both yeast and human RNA exosome complexes have revealed conservation in this structural organization of the RNA exosome (Figure 1B) [2, 5-8], suggesting evolutionary conservation not just within sequence but within structure and organization of the structural subunits.

The RNA exosome plays a pivotal role in processing, degradation, and surveillance of nearly every class of RNA in both the nucleus and cytoplasm [9-11]. First discovered as a crucial complex required for proper maturation of ribosomal RNA [3], the RNA exosome has since been shown to contribute to the processing of small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and transfer RNAs (tRNAs) [9, 12-16]. In addition, the RNA exosome is critical for RNA homeostasis within the nucleus

through targeting and degrading highly unstable species, such as cryptic unstable RNAs (CUTs) in *S. cerevisiae* and promoter upstream transcripts (PROMPTs) in human cells [14, 17-20]. The RNA exosome also plays a crucial role in RNA surveillance in both the nucleus and cytoplasm, degrading aberrant coding and non-coding RNA [20]. In addition to surveillance of misprocessed endogenous RNA species, the RNA exosome has been implicated in targeting foreign RNA through antiviral surveillance pathways [21].

Specificity of the RNA exosome for each target RNA is thought to be conferred through transient associations with nuclear and cytoplasmic cofactors [11]. The nuclear exoribonuclease Rrp6 and a stabilizing partner Rrp47 have been identified as cofactors with the budding yeast and human RNA exosome, which aid in post-transcriptional processing events of complex RNA such as pre-snoRNA, pre-snRNA, pre-tRNAs and 5.8S rRNA [2, 7, 22-24]. Structural studies show this interaction is facilitated by another nuclear cofactor, Mpp6, which provides a platform for direct cofactor interaction with the RNA exosome cap [24-28]. Together, these nuclear RNA exosome cofactors can recruit an essential nuclear DExD/H-box helicase, Mtr4 [23]. Mtr4 is a member of the larger interacting, and well characterized, TRAMP (Trf4/5-Air1/2-Mtr4 polyadenylation) complex. In budding yeast, the TRAMP complex has aids the RNA exosome in nuclear surveillance by targeting misprocessed rRNA, tRNA, snRNA, snoRNA as well as unstable CUTs for degradation [29-31]. Nuclear RNA surveillance in humans is also linked to a Mtr4-containing cofactor complex, the NEXT (nuclear exosome targeting) complex, which targets unstable PROMPTs in addition to misprocessed non-coding RNAs [32]. RNA exosome surveillance in the cytoplasm is also facilitated by interaction with a DExD/H-box helicase containing complex, the Ski complex, which assists in the degradation of aberrant mRNAs through decay pathways [22, 24, 33-35]. Though still not entirely understood and characterized, RNA exosome cofactor interactions tune the complexes ribonuclease activity, facilitating its roles in processing, degradation and/or surveillance events.

Though the RNA exosome is essential in all cell types and models tested thus far [3, 15, 36-38], recent clinical studies have identified pathogenic missense mutations in the structural subunit genes that result in distinct tissue-specific defects comprising a growing family of diseases termed RNA

exosomopathies [39]. Pathogenic missense mutations have been identified in the cap subunit genes *EXOSC1/2/3* and core subunit genes *EXOSC5/8/9* [40-51]. Missense mutations in cap subunit genes *EXOSC1/3* and in the structural core genes *EXOSC5/8/9* are linked to neurologic defects, with mutations in *EXOSC1/3/8/9* causing different forms of pontocerebellar hypoplasia and mutations in *EXOSC5* linked to cerebellar atrophy and SMA-like motor delays and hypotonia [40-42, 44-51]. Missense mutations in the cap subunit gene *EXOSC2*, however, cause a novel syndrome characterized by retinitis pigmentosa, hearing loss, premature aging and slight intellectual disability [43]. The diversity of RNA exosomopathy diseases continues to grow with new mutations identified by clinicians regularly. All the above listed RNA exosomopathy pathogenic amino acid substitutions and corresponding disease phenotypes, as well as some newly identified mutations, are summarized in Table S1.

While diverse in their clinical manifestations, these RNA exosomopathy missense mutations result in single amino acid substitutions in conserved domains of the structural subunits. Domain maps of *EXOSC2*, *EXOSC3*, *EXOSC5* and *EXOSC9* in Figure 1C are representative of the type of pathogenic amino acid substitutions reported in patients. RNA exosomopathy-linked mutations identified in the cap subunit genes *EXOSC2* and *EXOSC3* result in amino acid substitutions *EXOSC2* p.Gly198Asp (G198D) [43] and *EXOSC3* p.Asp132Ala (D132A) and p.Trp238Arg (W238R) [44, 45, 52, 53], respectively. These amino acid substitutions occur in not only highly conserved domains of both cap subunits as shown by the sequence alignments in Figure 1C. Furthermore, these pathogenic substitutions affect amino acids in similar regions of the two cap subunits, with *EXOSC2* Gly198 and *EXOSC3* Trp238 within or flanking a conserved structural “GxNG” motif within the RNA binding KH domain. The RNA exosomopathy-linked missense mutations identified in the core subunit genes *EXOSC5* and *EXOSC9* also result in amino acid substitutions in conserved domains of each protein. RNA exosomopathy mutations in *EXOSC5* result in amino acid substitutions p.Thr114Ile (T114I), p.Met148Thr (M148T), and p.Leu206His (L206H). These residues are located throughout the singular the PH domain of the core subunit, however the *EXOSC5* Leu206 falls closer the C-terminal end of the protein. Similarly, the *EXOSC9* RNA exosomopathy

pathogenic substitution falls near the termini of the protein. The RNA exosomopathy mutation in *EXOSC9* results in *EXOSC9* p.Leu14Pro (L14P), changing residue *EXOSC9* Leu14 to a proline near the N-terminus of the cores subunit.

Structural analysis of each RNA exosomopathy amino acid substitution suggest that these changes could affect inter-subunit binding interfaces or the conformation of the subunits themselves [6, 39, 47, 54, 55]. The residue *EXOSC2* Gly198 is positioned in a dense region of the subunit, surrounded by four β sheets and the substitution *EXOSC2* G198D is predicted to severely impact the structural organization of the cap subunit [43, 54]. The pathogenic substitutions in *EXOSC3* are predicted to impact interactions with surrounding subunits within the complex. The *EXOSC3* Asp132 residue is located in a loop between strands in the S1 domain and the substitution *EXOSC3* D132A likely would impair folding of the subunit and impact interactions with neighboring subunits *EXOSC5* and *EXOSC9* [39, 55]. Similarly, the *EXOSC3* Trp238 residue is predicted to position other *EXOSC3* residues to interact with neighboring *EXOSC9* residues, thus a substitution at this position could weaken *EXOSC3-EXOSC9* interactions [39, 55]. The *EXOSC5* pathogenic substitutions have been predicted to have destabilizing impacts on the subunit itself and complex interactions [47]. The *EXOSC5* Thr114 residue makes a hydrogen bond with another residue (Ala62) in the N-terminal of the subunit and the *EXOSC5* T114I substitution will likely disrupt this intra-subunit interaction [47]. The *EXOSC5* Met148 residue is at the interface with *EXOSC3* and a substitution to a polar Threonine likely will affect interactions between the subunits. The *EXOSC5* Leu206 residue is buried in a hydrophobic pocket of the subunit. Therefore, a substitution *EXOSC5* L206H is predicted to exert destabilizing effects on the integrity of the subunit [47]. Lastly, the *EXOSC9* Leu14 residue is located in the first alpha helix of *EXOSC5* and the substitution *EXOSC9* L14P may disrupt interactions within subunit [35]. Overall, these pathogenic amino acid substitutions are predicted to have varied biochemical structural consequences. As such, each RNA exosomopathy mutation may have differential impacts on the overall structure of the RNA exosome complex.

Several recent studies have begun investigating molecular consequences of these different pathogenic amino acid substitutions [55, 56]. Expression levels of EXOSC3 G31A and EXOSC3 W238R variants in a mouse neuronal line were reduced compared to wild type mouse EXOSC3, suggesting that these amino acid substitutions could affect the stability of the subunit [55]. Additionally, analysis of PCH patient fibroblasts and skeletal muscle cells homozygous for the *EXOSC9 L14P* mutations revealed that the EXOSC9 L14P levels are decreased compared to control samples, suggesting the pathogenic substitution impacts the stability of the subunit [42]. However, reconciling the diverse clinical pathologies seen in RNA exosomopathies is challenging if these different pathogenic amino acid substitutions simply decrease levels of the essential subunits and/or the level of the complex. Modeling these missense mutations and performing functional *in vivo* studies is critical to reveal the biology underlying the RNA exosomopathy diseases.

Recent studies modeling some of these RNA exosomopathy mutations in genetic model systems reveal distinct molecular and functional consequences resulting from the different pathogenic amino acid substitutions [47, 54-57]. These studies suggest that both complex integrity and interactions with known RNA exosome cofactors may be differentially impacted by specific RNA exosomopathy mutations [47, 54-57]. Any alteration in the RNA exosome levels or key cofactor interactions resulting from these amino acid substitutions would ultimately have an impact on the ability of the complex to process, degrade or survey RNA targets in the cell. Changes in RNA target levels could have a profound impact in certain tissues if key classes or specific RNAs are misprocessed, defective RNA accumulates and/or RNA homeostasis is dysregulated. While the previous *in vivo* studies of these RNA exosomopathy mutations provide valuable characterization *in vivo*, there has yet been a comparative assessment of the defects in RNA exosome function across multiple cap and core RNA exosomopathy mutant models. To comprehensively understand the distinct exosomopathy disease pathologies, it is necessary to comparatively assess how these exosomopathy amino acid substitutions affect the RNA exosome's ability in processing, degrading and surveying aberrant RNA *in vivo*.

To explore the functional consequences of these pathogenic amino acid substitutions within the RNA exosome, we took advantage of the budding yeast model system. Given that the RNA exosome was initially identified and has been most extensively studied in *Saccharomyces cerevisiae* [3, 58] and that the conservation in overall complex structure between the human and budding yeast RNA exosomes [28, 42, 59], a budding yeast system provides a robust platform to comparatively assess the *in vivo* consequences of exosomopathy mutations. We generated and analyzed *S. cerevisiae* models of the exosomopathy amino acid changes identified in EXOSC2, EXOSC3, EXOSC5, and EXOSC9 by mutating the corresponding budding yeast genes *RRP4*, *RRP40*, *RRP46*, and *RRP45*. We analyzed yeast cell growth and employed an unbiased RNA-Seq approach to explore the consequences of these missense mutations on the function of the RNA exosome. From these approaches, we detect the greatest functional defects in three of our mutant models, *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H*. The *rrp4-G226D* and *rrp40-W195R* cells model mutations *EXOSC2-G198D* and *EXOSC3-W238R* that are linked to SHRF and PCH, respectively [43, 46]. The *rrp46-L191H* cells model the mutation *EXOSC5-L206H* that causes cerebellar atrophy and neurological defects [47]. The *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells all show significant growth defects and have many differentially expressed genes, though to differing degrees. Comparative analysis of these differentially expressed genes across the three models show some shared changes, particularly in genes involved in rRNA processing and ribosome biogenesis, suggesting potential defects in translation within these RNA exosomopathy mutations. In addition, from our analysis we observe some differentially expressed genes that are unique to each of the three *rrp* mutant models, suggesting that while there are some shared consequences there are also distinct differences in RNA exosome function resulting from the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* mutations. Overall, these data represent an unbiased approach to comparatively characterize the *in vivo* defects in the function of the RNA exosome across a collection of RNA exosomopathy mutant models and suggests a new link between RNA exosomopathy defects and translation. This work also highlights the importance of *in vivo* functional studies to explore the consequences resulting from pathogenic amino acid changes that underlie different clinical presentations seen in RNA exosomopathy patients.

4.3 Materials and Methods

Chemicals and media

All chemicals were obtained from Sigma-Aldrich (St. Louis, MO), United States Biological (Swampscott, MA), or Fisher Scientific (Pittsburgh, PA) unless otherwise noted. All media were prepared by standard procedures [60].

Saccharomyces cerevisiae strains and plasmids

All DNA manipulations were performed according to standard procedures [61]. *S. cerevisiae* strains and plasmids used in this study are listed in Table S2. The *rrp4Δ*, *rrp40Δ*, *rrp45Δ* and *rrp46Δ* strains used in this study were previously described [62]. The *rrp6Δ* strain (ACY1641) was constructed by deletion of the *RRP6* open reading frame in the *BY4741* strain by homologous recombination using *RRP6-UTR natMX4* PCR products. The wild-type *RRP4 LEU* (pAC3656), *RRP40 LEU* (pAC3652), and *RRP46 LEU* (pAC3482) plasmids were constructed as previously described and contain *RRP4*, *RRP40* or *RRP46* endogenous promoter, terminator and 5'/3' UTR [47, 54, 55]. The *RRP45 LEU* (pAC3479) plasmid was generated by PCR amplification of the *RRP45* promoter and coding sequence from budding yeast genomic DNA using yeast gene-specific primers and cloning into pRS315[63] containing C-terminal 2xMyc tag and *ADHI* terminator. The 2xMyc-*ADHI* 3'-UTR was excised by restriction digestion and native *RRP45* 3' UTR was cloned into the *LEU* plasmid using NEBuilder HiFi Assembly (New England BioLabs). The different RNA exosomopathy amino acid substitutions were incorporated into the untagged *LEU2 RRP* expressing plasmids through site-directed mutagenesis with QuikChange Site-Directed Mutagenesis Kit (Stratagene) as previously described [47, 54, 55].

S. cerevisiae transformations and growth assays

All yeast transformations were performed according to the standard Lithium Acetate (LiOAc) protocol [64]. Cells were grown overnight to saturation in a 30°C incubator in liquid YEPD (1% yeast extract, 2% peptone, 2% dextrose, in distilled water). Cell concentrations were normalized to $OD_{600} = 0.4$ in 10 mL YEPD then incubated at 30°C for 5 hours. The cells were washed with TE/LiOAc then resuspended in TE/LiOAc to a concentration of 2×10^9 cells/mL. To these cells, plasmid DNA, single-stranded carrier DNA, and PEG/TE/LiOAc were added. The cells were agitated for 30 minutes at 30°C before adding DMSO. The cells were heat shocked at 42°C for 15 minutes, washed, and plated onto selective media.

Standard plasmid shuffle assays were performed to assess the *in vivo* function of the *rrp* variants as previously described [47, 54, 55]. The *rrp* Δ cells transformed with the *LEU RRP* wild-type control plasmid or the *LEU rrp* mutant variant plasmid were plated on 5-FOA Leu⁻ minimal media plates after 2-3 days of growth at 30°C. Single colonies from the 5-FOA Leu⁻ minimal media plates were selected in quadruplicate and streaked onto selective Leu⁻ minimal media plates. The cells containing only the *RRP/rrp LEU2* plasmid were used for the entirety of these studies.

The *in vivo* function of the *rrp* variants was assessed in growth assays on solid media and in liquid culture. The wild-type control *RRP* cells and the mutant model *rrp* cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation. Cell concentrations were normalized to an $OD_{600} = 0.5$, and samples were serially diluted in 10-fold dilutions and spotted onto Leu⁻ minimal media plates. Plates were grown at 25°C, 30°C, or 37°C for 2-3 days. For growth in liquid culture, cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation, diluted to an $OD_{600} = 0.01$ in Leu⁻ minimal media in a 24-well plate, and growth at 37°C was monitored and recorded at OD_{600} in a BioTek® SynergyMx microplate reader with Gen5™ v2.04 software over 24 hr. For both the liquid and solid media growth assay results shown, each sample was analyzed in at least 3 independent biological replicates. The liquid culture assays were performed in technical triplicate for each biological sample. Doubling times were calculated using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com), GraphPad Software, San Diego,

California USA.

Sample collection for RNA-Seq

RNA-Seq was performed on three independent biological replicates of *rrpΔ* cells containing the *RRP LEU* wild-type control plasmids or the *rrp LEU* variants as the sole copy of the RNA exosome gene. The *rrp6Δ* cells (ACY1641) contained either an *RRP6 LEU* wild-type control plasmid (pAC3752) or an empty *LEU* vector. Biological replicates of all samples were first screened by solid media growth assays prior to growth and collection for the RNA-Seq experiment. For sample collection, cells were grown in 2 mL Leu⁻ minimal media overnight at 30°C to saturation. Cell concentrations were normalized to an OD₆₀₀ = 0.2 in 15-30 ml of Leu⁻ minimal media and shifted to 37°C for 5 hours. Cells were washed, pelleted and flash frozen. Cell pellets were sent to Zymo Research and total RNA was extracted.

RNA-Seq Library Preparation

RNA-Seq library preparation was performed by Zymo Research. Total RNA-Seq libraries were constructed from 300 ng of total RNA. To remove rRNA, a method previously described [65] was followed with some modifications. Libraries were prepared using the Zymo-Seq RiboFree Total RNA Library Prep Kit (Cat # R3000) according to the manufacturer's instructions (<https://www.zymoresearch.com/products/zymo-seq-ribofree-total-rna-library-kit>). RNA-Seq libraries were sequenced on an Illumina NovaSeq to a sequencing depth of at least 20 million read pairs (150 bp paired-end sequencing) per sample.

Sequence Data Alignments and Differential Expression Analysis

NovaSeq paired-end 150-bp reads from Total RNA-Seq data files were first adaptor trimmed, and

then analyzed using the STAR program (version 2.6.1d) for alignment of short reads to *S. cerevisiae* reference genome. Transcript and gene expression estimates were measured using StringTie [v2.1.7 [66]]. The expression estimates fragments per Kilobase of transcript per Million mapped (FPKM) were used with the Pheatmap R package [v1.0.12 to generate heatmaps[67]]. The raw reads were per gene feature counted using featureCounts [v1.22.2 [68]] to the *S. cerevisiae* S288C genome assembly R64-1-1 [69], annotated with CUTs and SUTs [70]. Low feature counts (<10 reads) were removed. Differential gene expression analysis on raw read counts was performed using the DESeq2 R package [v1.38.1 [71]] to identify genes significantly changed (p -value<0.05, ≥ 1.5 -fold change) in *rrp* mutant variant samples relative to *RRP* wild-type control samples. Shrinkage of effect size was performed on differential expression data for visualizations using the *apeglm* method [72]. Using DESeq2, principal component analysis (PCA) was performed and MA plots were generated on raw read counts. Volcano plots of differential gene expression data were produced using EnhancedVolcano R package [v1.16.0.[73]]. UpSet plots were generated using UpSetR R package [v1.4.0[74]]. Piecharts and stacked bars of RNA class percentages in significantly altered genes were generated using GraphPad Prism version 9.3.1 for Windows (www.graphpad.com). Transcripts were classified by class using the annotations available through the *Saccharomyces cerevisiae* Genome Database (SGD) [75]. Gene Ontology (GO) analysis on significantly altered genes for Biological Process category was performed using the YeastMine webserver (<https://yeastmine.yeastgenome.org/yeastmine/begin.do>). GO analysis on human homolog genes was performed using HumanMine (<https://www.humanmine.org/humanmine>). All GO analyses were performed Holm-Bonferroni test correction. The full RNA-Seq datasets and GO analyses are compiled in Supplemental Documentation.

4.4 Results & Discussion

RNA exosomopathy mutations modeled in *Saccharomyces cerevisiae* orthologs have differential functional consequences.

To perform *in vivo* functional studies, we employed the budding yeast system to assess the differential consequences resulting from each RNA exosomopathy mutation. As shown in Figure 1C, the residues that are substituted in RNA exosomopathy patients are highly conserved, allowing for the variant to be readily modeled in *Saccharomyces cerevisiae*. We modeled the RNA exosomopathy mutations found in *EXOSC2/3/5/9* in the corresponding *S. cerevisiae* genes *RRP4/40/46/45* to generate budding yeast subunit variants containing the pathogenic amino acid substitutions. The SHRF-linked *EXOSC2 G198D* mutation is modeled in *rrp4-G226D* yeast cells that express the Rrp4 G226D cap subunit variant. The PCH-linked *EXOSC3-D132A*, *EXOSC3-W238R* and *EXOSC9-L14P* are modeled by the *rrp40-S87A*, *rrp40-W195R* and *rrp45-I15P* cells, which express the Rrp40 S87A, Rrp40 W195R or Rrp46 I15P variants, respectively. The *EXOSC5* RNA exosomopathy mutations *EXOSC5-T114I*, *EXOSC5-M148T* and *EXOSC5-L206H* are modeled by the *rrp46-Q86I*, *rrp46-L127T*, and *rrp46-L191H* cells, which express the Rrp46 Q86I, Rrp46 L127T and Rrp46 L191H variants. These corresponding *S. cerevisiae* RNA exosomopathy modeled amino acid substitutions are summarized in Table S1.

To assess functional consequences across these RNA exosomopathy mutant models, we performed a plasmid shuffle growth assay on solid minimal media in which cells were deleted for the genomic *RRP* gene and transformed with plasmids containing the different *rrp* alleles (See *Materials and Methods*). We also included a budding yeast strain (*BY4741*) within these experiments as a control for the genetic background of the *rrp4Δ*, *rrp40Δ*, *rrp45Δ* and *rrp46Δ* cells. Wild-type control cells *RRP4*, *RRP40*, *RRP45* and *RRP46* grow like that of the parental control strain at both temperatures tested (Figure 2A). Previous work has characterized the functional consequences of the *rrp4*, *rrp40* and *rrp46* mutant models [47, 54, 55]. In particular, these previous studies showed that the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* alleles can replace the essential *RRP* genes and that they each show growth defects compared to the corresponding wild-type control [47, 54, 55]. Consistent with these results, the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells show slower growth at 30°C and 37°C compared to corresponding wild-type controls (Figure 2A). Notably, the *rrp4-G226D* cells show the most severe growth defect at 37°C.

Furthermore, the *rrp40-S87A*, *rrp46-Q86I* and *rrp46-L127T* cells have no growth defects compared to the wild-type control *RRP40* and *RRP46* cells (Figure 2A). Similar to these mutant models, the *rrp45-I15P* cells show no difference in growth compared to *RRP45* wild-type control cells or the parental control cells (Figure 2A). We also included *rrp6Δ* cells which lack the RNA exosome cofactor Rrp6 as a comparative control for cells with disrupted RNA exosome function. The Rrp6 exonuclease is non-essential however the cofactor assists the RNA exosome in targeting and degradation of several key transcript RNAs [7, 22, 76, 77]. As expected from previous work [76], the *rrp6Δ* cells show extremely poor growth at 37°C compared to control cells. In comparing this slow growth phenotype to the slower growing RNA exosomopathy mutant models, the *rrp4-G226D* cells are most similar to the growth phenotype of *rrp6Δ*.

We also quantified to doubling time of this collection of RNA exosomopathy mutant model cells when grown at 30°C (Figure 2B) or 37°C (Figure 2C). We performed liquid growth assays and normalized all measurements to that of parental control cells at both temperatures. Consistent with the solid media growth assays, the *RRP* wild-type control cells have doubling times similar to the parental control cells at both 30°C and 37°C. Furthermore, the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells have significantly increased doubling times compared to control cells at both 30°C and 37°C. This quantification clearly reveals the difference in growth phenotypes between these three RNA exosomopathy mutant models. We observe that at 30°C, the *rrp46-L191H* cells have the longest doubling time compared to *rrp4-G226D* and *rrp40-W195R* cells, almost mirroring the doubling time measured for the *rrp6Δ* cells. However at 37°C we observe that the *rrp4-G226D* cells have the longest doubling time, comparable to that observed for the *rrp6Δ* cells.

Overall, these data suggest that RNA exosomopathy mutations have varied functional consequences *in vivo*. In particular, the SHRF-linked *EXOSC2-G198D* mutation, the PCH-linked *EXOSC3-W195R* mutation and the cerebellar atrophy-linked *EXOSC5-L206H* mutation, respectively, have the most profound functional consequences when modeled as the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* in the budding yeast system. These results are intriguing as the associated RNA exosomopathy disease

pathologies are diverse in the tissues impacted and the severity. The SHRF clinical outcome is relatively mild compared to the PCH symptoms associated with the *EXOSC5* mutations, and yet the *rrp4-G226D* mutant models show the most severe growth defect. Furthermore, these modeled mutations have been shown previously to have very little impact on the protein levels of the individual yeast Rrp subunits [47, 54, 55], suggesting that these growth defects are not simply due to varying levels of loss of the essential subunits and thus the complex. Taken together, and supported by previous evidence [47, 54, 55], these comparative growth data further suggest that the RNA exosomopathy pathogenic substitutions result in differential molecular consequences for the function of the RNA exosome *in vivo*. Using these *S. cerevisiae* models, we can comparatively assess the molecular consequences that may arise in the processing and/or degradation of RNA from these RNA exosomopathy mutations.

The Rrp4 G226D, Rrp40 W195R and Rrp46 L191H variants cause broad transcriptomic changes.

To unbiasedly investigate the molecular consequences of the different modeled RNA exosomopathy pathogenic amino acid substitutions, we performed RNA-Seq analysis on three independent biological replicates of the *rrp* mutant models and the corresponding wild-type *RRP* controls. Data collection and analysis was performed as described in *Materials and Methods* (full datasets available as Supplementary Document S1). The approaches taken to generate the *rrp* mutant models, as described in *Materials and Methods*, ensure that the genetic background for all mutants compared to corresponding wild-type controls is identical. Therefore, differences we observed in the *rrp* mutants from the RNA-Seq analysis can be attributed to the modeled pathogenic amino acid substitution.

Differential expression analysis was performed on each *rrp* mutant compared to its corresponding wild-type control [full lists of differential expressed genes ($\geq +1.5$ or ≤ -1.5 Fold Change [FC], $p < 0.05$) available as Supplementary Document S2]. This analysis reveals the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells have a large number of transcripts differentially increased or decreased compared to the corresponding *RRP4*, *RRP40* and *RRP46* control cells (Figure 3A). MA plots of the RNA-seq results for these three mutant variants further emphasize the large transcriptomic difference from the corresponding

wild-type control cells (Figure S1). The *rrp40-S87A* and *rrp46-L127T* have a number transcripts that are changed in comparison to *RRP40* or *RRP46* (Figure 3A), however there are fewer transcripts that are significantly increased or decreased as evidenced by the corresponding MA plots (Figure S1). The *rrp45-I15P* and *rrp46-Q86I* cells show very few transcriptomic differences compared to the corresponding *RRP45* or *RRP46* wild-type control cells (Figure 3A; Figure S1). Unbiased principal component analyses (PCA) of the RNA-Seq further indicates that that the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* mutant transcriptomes are distinct from those of the wild-type controls (Figure S1). Furthermore, the PCA analyses reveals reproducibility amongst the RNA-seq biological replicates (Figure S1), providing confidence in our differential expression analysis results. Visualizations of the differential expression analyses for the *rrp40-S87A*, *rrp45-I15P*, *rrp46-Q86I* and *rrp46-L127T* transcriptomes are presented in Figure S2.

From differential gene expression analysis of the *rrp4-G226D* cells, 516 transcripts are decreased (≥ -1.5 FC, $p < 0.05$) and 1196 transcript are increased ($\geq +1.5$ FC, $p < 0.05$) compared to the *RRP4* control (Figure 3B). Of the 516 transcripts decreased, a majority are mRNAs (86.96%) (Figure 3C), with the most significantly decreased transcript (≤ -3 FC) being *SSA1*, an mRNA that encodes a member of the Hsp70 chaperone family [78-80]. Interestingly, several other transcripts that encode heat shock protein family members are significantly decreased in *rrp4-G226D* cells, including *SSA2* (≤ -3 FC), which encodes an Hsp70 ATP-binding protein [80, 81], *HSC82* (≥ -3 FC), which encodes a cytoplasmic chaperone of the Hsp90 family [82, 83], and *HSP60* (≥ -2 FC) among others. Some of the most significantly decreased transcripts in *rrp4-G226D* cells are *RPS3* (≥ -2 FC) and *RPL15A* (≥ -2 FC), which encode protein components of the small and large ribosomal subunit, respectively [84]. The decrease of these ribosomal protein encoding transcripts is consistent with previous RNA-Seq analysis performed on *rrp4-G226D* cells [54]. Of the 1196 transcripts increased, only ~20% are mRNAs, with the majority being cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs), and other ncRNAs (Figure 3C). However intriguingly, the most significantly increased transcripts are two mRNAs, *PIR3* (≥ 3 FC) and *DDR2* (≥ 4 FC).

PIR3 encodes for an O-glycosylated cell wall protein that is required for cell wall stability [85] and *DDR2* encodes a multi-stress response protein [86].

From differential gene expression analysis of the *rrp40-W195R* cells, 426 transcripts are decreased (≥ -1.5 FC, $p < 0.05$) and 569 are transcripts increased ($\geq +1.5$ Fold Change [FC], $p < 0.05$) compared to the *RRP40* control (Figure 3D). Of the 426 transcripts decreased, a majority are mRNAs (81.92%) (Figure 3E), with the most significantly decreased transcripts being mRNAs involved in metabolic and biomolecular synthesis pathways including *URAI* and *URA4* (≤ -3 FC), which encode enzymes that catalyzes the steps in *de novo* synthesis of pyrimidines [87], *MDH2* (≤ -3 FC), which encodes a cytoplasmic malate dehydrogenase [88], and *HIS4* (≤ -4 FC), which encodes a multifunctional enzyme involved in histidine biosynthesis [89]. Among the most significantly decreased transcripts in *rrp40-W195R* cells are mRNAs *RPS13* (≥ -2 FC) and *RPS7B* (≥ -2 FC) which encode protein components of the small and large ribosomal subunit, respectively [84]. Of the 569 transcripts increased, a majority are ncRNAs such as CUTs (29.15%), SUTs (15.02%), snoRNAs (5.30%) and tRNAs (5.83%) (Figure 3E). Many of the most significantly increased transcripts in *rrp40-W195R* cells are uncharacterized ORFs and ncRNAs, with snoRNAs *snR66* and *snR65* among the most significantly increased transcripts detected (≥ 3 FC). As observed in the *rrp4-G226D* transcriptome, only ~20% of transcripts significantly increased in *rrp40-W195R* cells are mRNAs. However one of the most significantly increased transcripts in *rrp40-W195R* cells is *RRN3* which encodes an essential transcription factor for RNA polymerase I [90].

From differential gene expression analysis of the *rrp46-L191H* cells, 487 transcripts are decreased (≥ -1.5 FC, $p < 0.05$) and 724 transcripts are increased ($\geq +1.5$ Fold Change [FC], $p < 0.05$) compared to the *RRP46* control (Figure 3F). Of the 487 transcripts decreased, ~60% are mRNAs while the other are ncRNAs, with nearly 10% being tRNAs, and a combined ~20% being CUTs, SUTs and uncharacterized/dubious ORFs (Figure 3G). However, the most significantly decreased transcripts in *rrp46-L191H* cells are mRNAs involved in biomolecular synthesis pathways, including *HIS4* (≤ -4 FC), which encodes a multifunctional enzyme involved in histidine biosynthesis [89], *URAI* and *URA4* (≤ -2 FC), which

encode enzymes that catalyze steps of *de novo* pyrimidine biosynthesis [87], *BIO3* and *BIO4* (≤ -4 FC), which encode a synthetase and an aminotransferase involved in biotin synthesis, respectively, [91, 92], and, lastly, *RIB4* (≤ -3 FC), which encodes a synthase that catalyzes the synthesis of a riboflavin precursor [93]. As observed for the *rrp4-G226D* cells, two of the most significantly decreased transcripts in *rrp46-L191H* cells are ribosomal subunit mRNAs *RPS3* and *RPL15A* (≥ -2 FC). Of the 724 transcripts increased, $\sim 30\%$ are classified as CUTs, $\sim 15\%$ classified as SUTs and $\sim 30\%$ being mRNAs (Figure 3G). The most significantly increased transcripts however are two mRNAs *RPL18B* (≥ 3 FC), which encodes a component of the large ribosomal subunit [84], and *CBT1* (≥ 4 FC) which encodes a protein involved in 5' RNA end processing of mitochondrial cytochrome *b* mRNA [94]. Intriguingly, *CBT1* is also linked to processing of 15S rRNA [95].

Overall, these differential expression analyses reveal many broad transcriptomic changes that can partially explain the growth defects observed in the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells. The significant decrease in several transcripts that encode various Hsp family members in the *rrp4-G226D* cells could suggest that these cells have compromised response to heat stress, thus explaining the significant growth defect observed at 37°C (Figure 2). Additionally, previous work has shown that a loss of Rrp6 lead to decreased expression of *HSP* genes in a manner independent from the interaction between the exonuclease and the RNA exosome complex [96]. Perhaps the Rrp4 G226D variant could impact this moonlighting role of Rrp6 *in vivo*. This compromised response to stress may also explain the increase in *DDR2* and *PIR3* mRNA transcripts observed in the *rrp4-G226D* cells as expression of both these genes is activated in response to a variety of stressful conditions [97, 98]. Additionally, the decrease in several biosynthetic transcripts observed in the *rrp40-W195R* and *rrp46-L191H* could also reflect overall slowed growth. Intriguingly, both *rrp40-W195R* and *rrp46-L191H* cells share similar decreases in *URA1* and *URA4* transcripts, perhaps suggesting a link between the Rrp mutant variants and pyrimidine levels *in vivo*. Intriguingly, all three mutant variants show decrease in mRNAs that encode components of the ribosome

within the top significantly decreased transcripts, suggesting these RNA exosomopathy mutations could affect ribosome biogenesis.

These broad transcriptomic changes also reflect defects in the function of the RNA exosome. Consistent with the role the RNA exosome plays in degradation of nascent ncRNA species, the CUTs and SUTs combined make up the largest group of transcripts significantly increased in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells (Figure 3). As CUTs and SUTs are stabilized in RNA exosome mutants and cross-link to the RNA exosome [14, 18, 99, 100], these transcripts are likely direct targets of the RNA exosome and the observed increase is indicative of disrupted function of the complex. This significant increase in CUTs and SUTs is also consistent with previous RNA-Seq analysis of the *rrp4-G226D* cells [54].

Comparison across the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* mutant models is intriguing when we consider the overall classes of RNAs that are significantly increased or decreased (Figure 3C, 3E and 3G). As illustrated in Figure 3, a majority of transcripts decreased in all three mutants are mRNAs. However in the *rrp46-L191H* mutant, only 58% of the significantly decreased transcripts are mRNAs, as compared to 80-90% in both the *rrp4-G226D* and *rrp40-W195R* cells. In the *rrp46-L191H* cells, nearly 10% of decreased transcripts are tRNAs and 10% are CUTs and SUTs. In the *rrp4-G226D* cells, ~5% of decreased transcripts are tRNAs, but the level of decreased CUTs and SUTs is much lower. This decrease in tRNAs is not shared in the *rrp40-W195R* cells, which only show 0.23% of decreased transcripts that are tRNAs. These divergent changes are even more apparent when we compare the comparing categories of significantly increased transcripts. While there is an increase in CUTs and SUTs across all three mutant models, these non-coding transcripts comprise different percentages of the transcripts affected in each *rrp* mutant. In particular, the *rrp4-G226D* cells have the largest percentage of SUTs increased the three mutants. There are also distinct increases in other ncRNAs across the three mutants, including snoRNAs and tRNAs. The *rrp40-W195R* cells show the highest categorization of differentially increased snoRNAs (5.3%) and tRNAs (5.8%) among the three mutants. Overall the differential expression data suggest that the *rrp4-*

G226D, *rrp40-W195R* and the *rrp46-L191H* cells have distinct transcriptomic changes as compared to corresponding wild-type control cells. Furthermore these changes also differ among the three mutants, suggesting differential consequences resulting from the Rrp variants *in vivo*.

Comparative assessment of differentially expressed transcripts within *rrp4-G226D*, *rrp40-W195R* and the *rrp46-L191H* suggests shared impacts on metabolic pathways and rRNA modification and processing.

To further comparatively assess the molecular consequences resulting from the modeled RNA exosomopathy mutations in these yeast mutant models, we investigated what decreased and increased transcripts are shared across the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* models. To do so, we generated UpSet plots on lists of differentially expressed transcripts that were decreased by ≤ -1.5 Fold Change (FC) or increased by $\geq +1.5$ Fold Change (FC) significantly ($p < 0.05$) within the three mutant models to identify shared impacts (Figure 4). From the UpSet plot, we detect 86 transcripts shared that are significantly decreased ≤ -1.5 FC (Figure 4A) and 209 shared transcripts that are significantly increased $\geq +1.5$ (Figure 4B) among the *rrp4-G226D*, *rrp40-W195R* and the *rrp46-L191H* cells (Figure 4A). Of these 86 decreased transcripts, 89.5% encode for mRNAs (Figure 4B). Gene Ontology (GO) analysis on these 86 shared decreased transcripts reveal metabolic and biosynthetic biological processes are the top significantly impacted, with carboxylic acid (GO:0019752), oxoacid metabolic (GO:0043436), and organic acid metabolic process (GO:0006082) the most significantly enriched processes (Figure 4C). This suggests that the *rrp4-G226D*, *rrp40-W195R* and the *rrp46-L191H* cells all have significantly decreased transcripts that impact metabolic and biosynthetic pathways. These data align with the transcripts that are most significantly decreased in the *rrp40-W195R* and *rrp46-L191H* cells. Of the 209 transcripts increased, a majority are CUTs and SUTs, consistent with the trend observed for each *rrp* mutant model (Figure 4E). GO analysis on these increased transcripts reveals however that rRNA modification (GO:0000154) is the most significant biological process enriched (Figure 4F). This GO analysis result is likely due to the 5.8% of

increased transcripts that are snoRNAs. Surprisingly, sexual sporulation (GO:0034293) and meiotic cell cycle (GO:0051321) are two significantly enriched biological processes within these shared increased transcripts. The enrichment of these two biological processes is likely due to the shared mRNA transcripts that are increased in *rrp4-G226D*, *rrp40-W195R* and the *rrp46-L191H* cells, as GO analyses of the 12.62% mRNAs within the shared increased transcripts showed meiotic cell cycle (GO:0051321) as the most significantly enriched process (Supplemental Documentation S3). However, the significant enrichment in rRNA modification, as well as in rRNA methylation (GO:00031167) and rRNA processing (GO:0006364) that also emerges from the GO analysis (Figure 4F), suggests that there could be shared impacts on ribosome biogenesis within these three *rrp* mutant models.

To further assess shared changes, we performed GO analyses on the identifiable human homologs of the shared 86 decreased transcripts and 209 increased transcripts. The same GO biological processes most significantly enriched from the budding yeast transcripts, carboxylic acid (GO:0019752), oxoacid metabolic (GO:0043436), and organic acid metabolic process (GO:0006082), are the most significantly enriched processes for the human homologs (Supplementary Documentation S4). This suggests that these modeled RNA exosomopathy mutations result in changes in highly conserved metabolic and biosynthetic pathways. Of the shared 206 increased transcripts, few had identifiable human homologs as a majority are yeast-specific CUTs and SUTs. However, GO analysis of those transcripts that did have identifiable human homologs reveals a significant enrichment in synaptic vesicle priming and fusion biological processes (GO:0016082; GO:0031629; GO:0099500). The link between these modeled RNA exosomopathy mutations and biological processes involved in synaptic vesicle fusion and trafficking as demonstrated by the GO analysis of the human homologs of increased transcripts identified within our RNA-Seq experiment could provide context for the numerous neurological defects that are common in RNA exosomopathy patients. Additionally, the changes in conserved metabolic pathways we observed in our GO analyses of the decreased transcripts could also contributed to the disease pathologies in these individuals with RNA exosomopathies.

The UpSet plots in Figure 4 can also be employed to identify transcripts that are shared between two of the three *rrp* mutants (Figure 5). We identify 97 shared transcripts that are significantly decreased in both *rrp40-W195R* and *rrp46-L191H* cells, 100 shared transcripts that are significantly decreased in both *rrp40-W195R* and *rrp4-G226D* cells, and 50 shared transcripts that are significantly decreased in *rrp4-G226D* and *rrp46-L191H* cells (Figure 5A). A majority of the decreased transcripts shared between *rrp40-W195R* and *rrp46-L191H* and between *rrp40-W195R* and *rrp4-G226D* are mRNAs (Figure 5B). Intriguingly, a large percentage of shared tRNAs that are significantly decreased in the *rrp4-G226D* and *rrp46-L191H* cells (Figure 5B). GO analyses on these sets of shared transcripts reveal enrichment in different biological processes related to translation. A significant number of the 100 shared transcripts significantly decreased in both *rrp40-W195R* and *rrp4-G226D* cells impact cytoplasmic translation (GO:0002181) (Figure 5C). A significant number of the 50 shared transcripts significantly decreased in both *rrp4-G226D* and *rrp46-L191H* cells impact translation elongation (GO:0006414) (Figure 5D), consistent with the large percentage of tRNAs that are decreased in both of these models. No significant enrichment of any specific biological process is detected for the 97 shared transcripts significantly decreased in the *rrp40-W195R* and *rrp46-L191H* cells. Both GO analyses of the shared increased transcripts between *rrp40-W195R* and *rrp4-G226D* and *rrp40-W195R* and *rrp46-L191H* reveal significant enrichment of biosynthetic processes, such as amide biosynthesis (GO:0043604), organonitrogen compound biosynthesis (GO:1901566) and cellular macromolecule biosynthesis (GO:0034645).

The UpSet plots can also be employed to identify 54 shared transcripts that are significantly increased in both *rrp40-W195R* and *rrp46-L191H* cells, 152 shared transcripts that are significantly increased in both *rrp40-W195R* and *rrp4-G226D* cells, and 268 shared transcripts that are significantly increased in *rrp4-G226D* and *rrp46-L191H* cells (Figure 5E). A majority of increased transcripts shared between *rrp40-W195R* and *rrp46-L191H* are mRNAs (Figure 5F). GO analysis on these 54 transcripts reveal a significant enrichment in rRNA metabolic process (GO:0016072), ncRNA processing (GO:0034470), rRNA processing (GO:0006364) and ribosome biogenesis (GO:0042254) (Figure 5G).

Classification of the 152 shared transcripts that are significantly increased in both *rrp40-W195R* and *rrp4-G226D* cells reveals that only a proportion of those changed are mRNAs, with a larger majority non-coding CUTs and SUTs (Figure 5F). GO analysis of these 153 shared transcripts reveal no significant enrichment of any biological process. Classification of the 268 shared transcripts that are significantly increased in *rrp4-G226D* and *rrp46-L191H* cells also reveals a majority are CUTs and SUTs. In contrast to the shared transcripts between *rrp40-W195R* and *rrp4-G226D*, GO analysis of these 268 transcripts reveal significant enrichment of telomere maintenance (GO:0000722) and mitotic recombination (GO:0006312). The enrichment in telomere maintenance and mitotic recombination is likely due the shared increase in *YRF* transcripts which encode several telomeric Y' element DNA helicases [101]. Increased *YRF* levels indicate compensation for telomeric loss [101]. This link between *rrp4-G226D* and *rrp46-L191H* cells and telomere health is of interest as the RNA exosome processes the telomerase component RNA, *TLC1* [102], and previous studies have shown significant increases in steady state levels of the *TLC1* precursor in both the *rrp4-G226D* and *rrp46-L191H* cells [47, 54]. Overall these analyses of the shared targets between the pairs of RNA exosomopathy mutant models—*rrp40-W195R* and *rrp46-L191H*, *rrp40-W195R* and *rrp4-G226D*, and *rrp4-G226D* and *rrp46-L191H*—can reveal more insight into molecular consequences resulting from the Rrp variants. These results provide some intriguing connections between biological pathways and these different RNA exosomopathy mutations. Moreover, these analyses reveal that different processes are not specifically impacted by the type of RNA exosome subunit that is mutated, as the *rrp40-W195R* and *rrp4-G226D* cells, modeling mutations in the cap subunit genes, show distinctly different shared targets with the *rrp46-L191H* cells, modeling a mutation in a core subunit gene.

Comparative assessment of differentially expressed transcripts specific to *rrp4-G226D*, *rrp40-W195R* or *rrp46-L191H* suggest distinct impacts on translation and ribosome biogenesis.

Transcripts that are significantly increased or decreased specifically in each of the *rrp* mutant cells can also be examined through the UpSet plots (Figure 6). There are 143 transcripts that are significantly decreased

specifically in the *rrp40-W195R* cells, 254 transcripts significantly decreased specifically in the *rrp46-L191H* cells and 280 transcripts significantly decreased specifically in the *rrp4-G226D* cells (Figure 6A). From this analysis, distinct patterns emerge comparing *rrp46-L191H* and the cap mutant models (Figure 6B). In particular, a majority of transcripts decreased specifically in either the cap mutant models are mRNAs, while a plurality of transcript types are impacted in the *rrp46-L191H* cells with about a third comprised of tRNAs, CUTs and SUTs (Figure 6B). GO analyses of these decreased transcripts specific to each *rrp* mutant reveal similarity, however, between the *rrp46-L191H* and *rrp4-G226D* decreased transcripts (Figure 6C-E). GO analysis of the 143 transcripts that are decreased specifically in the *rrp40-W195R* reveal significant enrichment in biological processes related to cytoplasmic translation (GO:0002181) and ribosomal small subunit assembly (GO:0000028) (Figure 6C). GO analyses of the 54 transcripts decreased specifically in the *rrp46-L191H* cells and the 280 transcripts decreased specifically in the *rrp4-G226D* cells also show significant enrichment in processes related to translation (GO:0006412) (Figure 6D-E). However, these analyses also reveal significant enrichment in biosynthetic processes, particularly amide and peptide biosynthesis and metabolic processes (Figure 6D-E).

We can also identify transcripts that are significantly increased specifically in each of our *rrp* mutants. We identify 154 transcripts that are significantly increased specifically in the *rrp40-W195R* cells, 193 transcripts significantly increased only in the *rrp46-L191H* cells and 567 transcripts significantly increased only in the *rrp4-G226D* cells (Figure 6F). Analysis of these individual sets of changes reveals a divergent pattern between the three *rrp* mutant models (Figure 6G). Of the 154 transcripts that are increased specifically in the *rrp40-W195R* cells, nearly a quarter comprise snoRNAs, snRNAs, and tRNA. Another quarter of these 154 transcripts are CUTs and SUTs, a third quarter are mRNAs, and the fourth quarter mostly dubious or uncharacterized open reading frames (ORFs). Intriguingly, the transcripts increased only in *rrp46-L191H* cells show a different pattern, with a large majority of those 193 RNAs consisting of mRNAs. Lastly, the *rrp4-G226D*-specific increased RNAs show yet a different pattern, with a majority being CUTs and SUTs. GO analysis of the 154 transcripts significantly increased in only the *rrp40-W195R*

cells reveals significant enrichment in biological processes involved in gene expression (GO:0010467), rRNA modification (GO:00000154), and translation elongation (GO: 0006414). GO analysis of the 193 transcripts increased specifically in the *rrp46-L191H* cells reveal significant enrichment in processes related to ncRNA processing (GO:0034470) and ribosome biogenesis (GO:0042254). GO analysis of the transcripts increased only in the *rrp4-G226D* cells reveals no significant enrichment, likely due to the large percentage of CUTs and SUTs.

Overall though, this analysis gives us further insight into the molecular consequences resulting from each RNA exosomopathy mutation modeled in the budding yeast system. The GO analyses performed on the transcripts changed solely in the *rrp* mutants did reveal enrichment in several biological processes that were identified in other GO analyses already discussed (Figure 4-5). However the transcripts used to produce the GO terms in Figure 6 are those significantly changed specifically within each cell type—*rrp4-G226D*, *rrp40-W195R* or *rrp46-L191H*. Therefore, while there may be overlapping impacts on key biological processes within these cells, these consequences are in part due to distinct targets.

Heatmaps of FPKMs reveal similar pattern in overall *RSP*, *RPL*, *CUTs* and *SUTs* levels between RNA exosomopathy mutant models and *rrp6Δ* cells.

Within all of the GO analyses, common GO terms enriched are related to translation and ribosome biogenesis. Additionally, within the differential expression analysis, some of the most significantly decreased transcripts in *rrp4-G226D*, *rrp40-W195R* or *rrp46-L191H* were *RPS* and *RPL* mRNAs which encode components of the ribosome [84]. The ribosome produces all cellular proteins and is a large complex containing numerous conserved components [103, 104]. The eukaryotic ribosome consists of a small (40S) and large (60S) RNA-protein subunits. In *Saccharomyces cerevisiae*, the 40S subunit consists of 33 ribosomal proteins (Rps prefix) and the 18S ribosomal RNA (rRNA), while the 60S subunit (60S) consists of 46 ribosomal proteins (Rpl prefix) and 3 rRNAs (25S, 5.8S and 5S rRNA). Notably, decreases in *RPS*

and *PRL* mRNAs have also been observed in *rrp6Δ* mutant cells [105]. To broadly compare impacts on ribosomal protein genes across the *rp4-G226D*, *rrp40-W195R* and *rrp46-L191H* samples, we generated heatmaps of normalized FPKM expression estimates specifically for the *RPS* and *RPL* genes (Figure 7A). We included FPKM estimates for *rrp6Δ* cells that were collected in the same RNA-Seq experiment. Consistent with previous work, there is a broad decrease in expression levels for most ribosomal protein genes within *rrp6Δ* samples. We also detect an overall decrease in ribosomal protein gene expression within the three RNA exosomopathy models with the triplicates clustering together, though the decrease is less than the decrease observed in the *rrp6Δ* samples. However, the *rrp4-G226D* samples show the broadest decrease in expression of ribosomal protein transcripts compared to the *rrp40-W195R* and *rrp46-L191H* samples, with the three triplicates clustering together, separate from the other *rrp* mutant samples. The *rrp40-W195R* samples show the fewest change in expression levels of *RSP* and *RPL* transcripts, with one of the triplicates clustering with the wild-type controls. A larger heatmap with gene identifiers is shown in Figure S3.

A large increase in CUTs and SUTs is detected within all three *rrp* mutant models. Currently there is no GO annotation associated with budding yeast CUTs and SUTs, therefore the GO analyses do not include these transcripts. Therefore, we generated a heatmap of normalized FPKM expression estimates to compare the impact these *rrp* mutations have on the CUTs and SUTs (Figure 7B). We included FPKM estimates of the *rrp6Δ* samples as these transcripts were first identified by deletion of *RRP6* and Rrp6 activity is important for degradation of the CUTs [70, 99, 106]. As supported by these previous data, we see a substantial increase in expression of CUTs and SUTs within the *rrp6Δ* samples. We also see an increase in CUTs and SUTs within the RNA exosomopathy mutant models, with *rp4-G226D*, *rrp40-W195R* and *rrp46-L191H* clustering together. The *rrp4-G226D* cells show the most increase in CUTs and SUTs, again with the triplicates clustering together, separately from the other *rrp* mutant samples. Similar to the *rrp4-G226D* samples, *rrp40-W195R* and *rrp46-L191H* do show some increased expression of CUTs and SUTs, with the triplicates clustering together.

While the heatmaps presented in Figure 7 show similar broad decreases both for ribosomal protein transcripts and increases in CUTs and SUTs across the RNA exosomopathy mutant models, they also show that not all the same transcripts are changed in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells. This is particularly evident in the heatmap comparing CUTs and SUTs gene expression (Figure 7B). The *rrp6Δ* samples show a broad, indiscriminate increase in all CUTs and SUTs. This result is likely due to the fact that these transcripts were initially identified based upon their accumulation in *rrp6Δ* cells [70]. Regardless, this increase can be interpreted as disrupted RNA exosome targeting and function. While the *rrp4-G226D* cells show the broadest increase in CUTs and SUTs of the three RNA exosomopathy mutant models assessed, there are definitive groups of CUTs and SUTs that appear to not be increased. Similarly, *rrp40-W195R* and *rrp46-L191H* show some shared and some distinct changes in CUTs and SUTs expression. These observations suggest that within all three RNA exosomopathy mutant models the RNA exosome targeting and degradation of these transcripts is impacted, yet in distinct ways as only specific targets are elevated. Furthermore, CUTs are transcriptionally terminated and targeted to the RNA exosome via the Nrd1-Nab3-Sen1 (NNS) transcription termination complex [107-109]. Accumulation of these transcripts in the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells may suggest consequences for NNS complex interactions and disrupted nuclear surveillance of pervasive transcription.

4.5 Conclusion

The work presented here represents the first *in vivo* comparative study of this collection of RNA exosomopathy mutant models. The RNA-Seq experiment provides many intriguing results and implicates metabolic and biosynthetic processes, rRNA processing/ modifications and ribosome biogenesis as potential pathways impacted within these *rrp* mutant cells. Future studies are needed to mechanistically understand whether these molecular consequences are directly caused by the Rrp variant or are secondary consequences resulting from defects in the function of the RNA exosome. Furthermore, biochemical studies can shed light on if the pathogenic amino acid substitutions impact the function of the complex by altering complex integrity or disrupting key cofactor interactions. Future studies can also take advantage of gene

editing technology, such as CRISPR/Cas9, to more precisely assess direct causes resulting from the RNA exosomopathy and alleviate some bias that can be introduced with tradition genetic and biochemical techniques (a potential methodology to edit *RRP* genes to express RNA exosomopathy mutations is outlined in Appendix I). However regardless, the information captured in the RNA-Seq data and analysis provides a platform for future research studies and avenues of investigation that will further our understanding of the *in vivo* consequences resulting from these different RNA exosomopathy mutations.

4.6 Chapter IV Figures

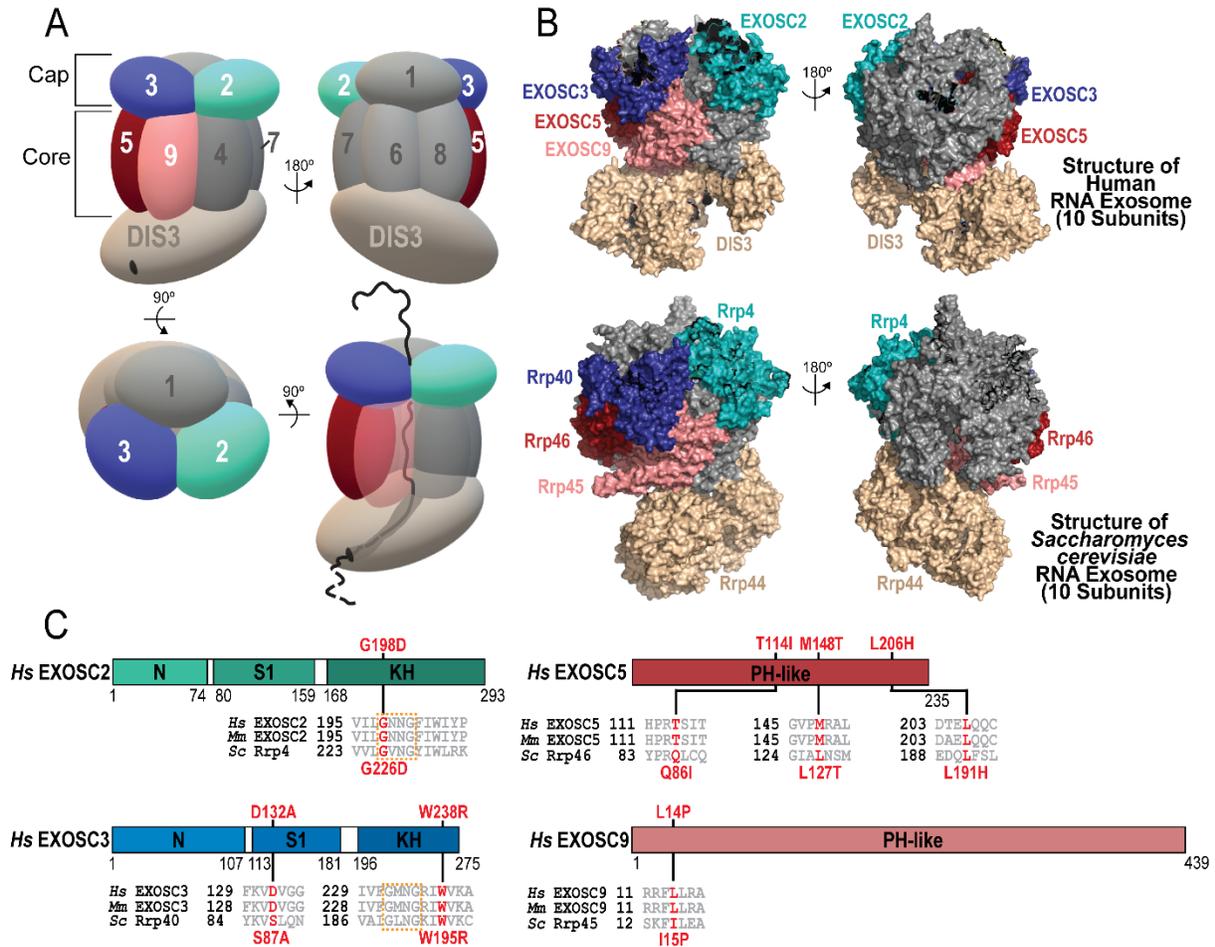


Figure 1. Overview of pathogenic amino acid substitutions in the human cap and core structural subunits of the RNA exosome.

(A) The RNA exosome is an evolutionary conserved ribonuclease complex composed of nine structural subunits (EXOSC1-9) and one catalytic subunit (DIS3) that form a “cap” and “core” ring-like structure. The 3-subunit cap is composed of EXOSC1/Csl4 (Human/*S. cerevisiae*), EXOSC2/Rrp4, and EXOSC3/Rrp40. The 6-subunit core is composed of EXOSC4/Rrp41, EXOSC5/Rrp46, EXOSC6/Mtr3, EXOSC7/Rrp42, EXOSC8/Rrp43, and EXOSC9/Rrp45. The DIS3/Dis3/Rrp44 catalytic subunit is located at the bottom. The structural cap and core subunits form a barrel-like structure through which RNA is

threaded in a 5'-3' orientation. Recently, missense mutations linked to novel RNA exosomopathy diseases have been identified in a number of these structural *EXOSC* genes. Missense mutations in the gene encoding the EXOSC2 cap subunit (teal blue, labeled 2,) are linked to a novel syndrome termed SHRF (short stature, hearing loss, retinitis pigmentosa and distinctive facies) [43]. In contrast, missense mutations in the gene encoding the EXOSC3 cap subunit (dark blue, labeled 3) and EXOSC9 core subunit (pink, labeled 9) cause forms of PCH (pontocerebellar hypoplasia), a severe disease characterized by early onset atrophy of the pons and cerebellum [40, 42, 44-46, 50, 51, 110]. Missense mutations in the gene encoding the EXOSC5 core subunit (red, labeled 5) are linked to a disease characterized by cerebellar atrophy, SMA-like motor delays and hypotonia [47]. These differing RNA exosomopathy missense mutations result in single base changes within each subunit. (B) The structure and organization of the RNA exosome is highly conserved across eukaryotes. A structural model of the human RNA exosome (left) [PDB 6D6Q] [23] and the *S. cerevisiae* RNA exosome (right) [PDB 6FS7] [111] are depicted with the core and cap subunits that are linked to RNA exosomopathy diseases labeled and color coded. The human cap subunits EXOSC2 and EXOSC3 correspond to *S. cerevisiae* orthologs Rrp4 and Rrp40. The human core subunits EXOSC5 and EXOSC9 correspond to *S. cerevisiae* orthologs Rrp46 and Rrp45. (C) Domain maps are shown for EXOSC2, EXOSC3, EXOSC5 and EXOSC9. Both cap subunits are composed of three different domains: an N-terminal domain, a central putative RNA binding S1 domain, and a C-terminal putative RNA binding K homology (KH) domain. The "GxNG" motif identified in the KH domain of both cap subunits is boxed in orange. The GxNG motif may play a structural role as it is buried at the interface between the S1 and KH domains in the 3D structure of the cap subunits [112]. Both cores subunits are composed of a singular PH-like domain. The position of the RNA exosomopathy disease-linked amino acid substitutions in the human subunits are depicted above the domain structures in red. Sequence alignments of EXOSC/Rrp orthologs from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*) and *S. cerevisiae* (*Sc*) below the domain structures reveal the high conservation of the residues altered in disease (in red) and the sequences flanking these residues (in gray). The amino acid substitutions generated in the Rrp orthologs for this study that correspond to the pathogenic amino acid substitutions are shown below the sequence alignments in red.

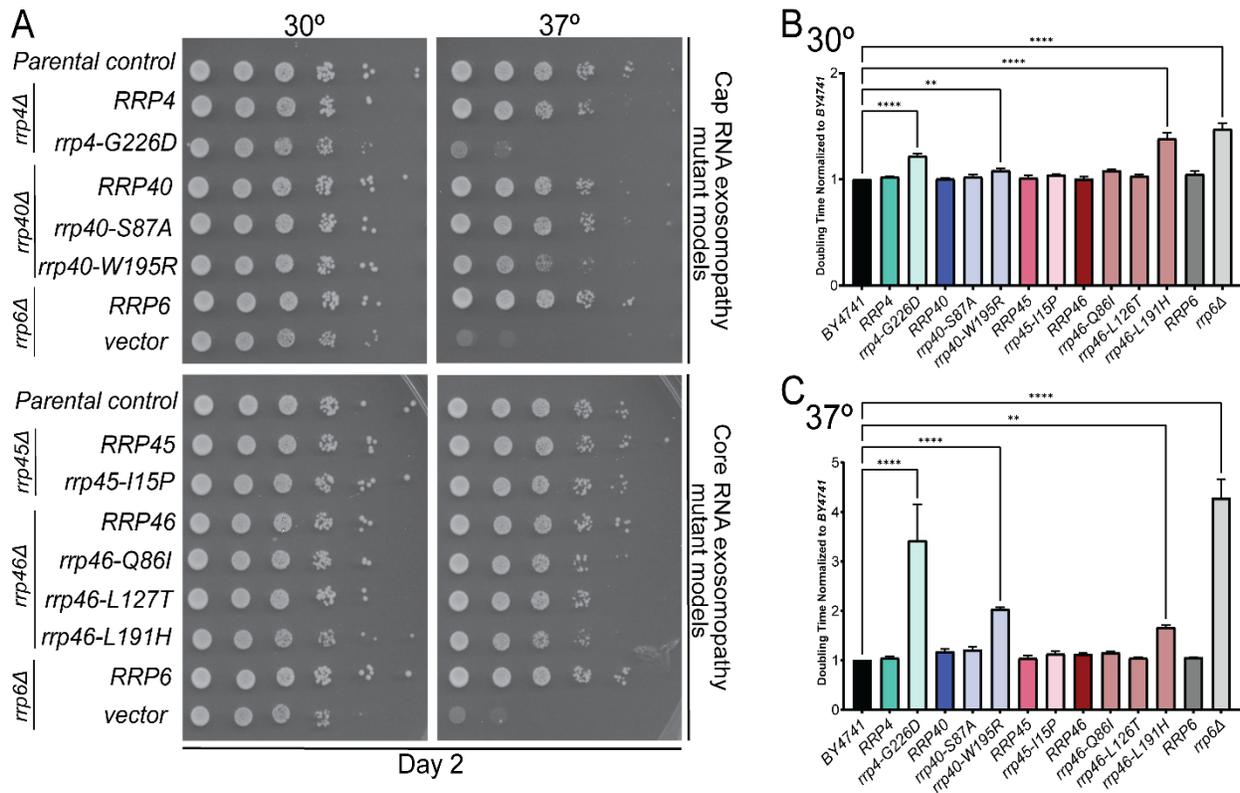


Figure 2. *S. cerevisiae* Rrp4 variants that model EXOSC variants identified in patients show impaired function.

S. cerevisiae cells expressing Rrp variants that model pathogenic amino acid changes found in EXOSC2, EXOSC3, EXOSC5 and EXOSC9 were generated as described in *Materials and Methods*. Growth assays reveal functional consequences of different *S. cerevisiae* *rrp* mutants that are modeling pathogenic amino acid substitutions identified in RNA exosome cap and RNA exosome core subunits. The growth of *rrpΔ* cells expressing a *rrp* mutant variant or the corresponding wild-type *RRP* control were analyzed (A) by serially diluted, spotted onto solid media grown at the indicated temperatures and (B-C) by liquid growth to quantify cell doubling times. Included in the assays as comparative controls are *BY4741* cells expressing an empty vector (labeled “Parental control”) and *rrp6Δ* cells, expressing *RRP6* or an empty vector. (A) Solid media growth assays reveal that the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells exhibit

impaired growth compared to control *RRP4*, *RRP40*, *RRP46* and *BY4741* cells. The *rrp4-G226D* cells reveal the most profound growth defect at 30°C and 37°C, nearly phenocopying the impaired growth observed in the comparative control *rrp6Δ* cells expressing an empty vector at the same temperatures. The *rrp40-W195R* cells show impaired growth at 37°C though appear to grow similarly to *RRP40* cells on the solid media at 30°C. The *rrp40-S87A* cells grow similarly to the *RRP40* and *BY4741* cells at both temperatures. The *rrp46-L191H* cells show impaired growth at 37°C and slightly slower growth at 30°C compared to the *RRP46* wild-type control. However the other *rrp46* variants do not show any impaired growth at either temperature compared to the *RRP46* wild-type controls or the *BY4741* cells. The *rrp45-I15P* mutant cells show no difference in growth compared to the wild-type control *RRP45* and *BY4741* cells. (B-C) Liquid growth measurements of each sample in technical triplicate were recorded for cells at both 30°C and 37°C. Quantified doubling times for each sample were normalized to the value of the *BY4741* samples at 30°C and 37°C. Consistent with the observed growth on the solid media plates, all *RRP* wild-type control cells have doubling times nearly equivalent to the *BY4741* cells. The *rrp4-G226D* cells show significantly increased doubling time compared to the *BY4741* control at both temperatures tested. Additionally the *rrp40-W195R* cells show significantly increased doubling time at both temperatures, though the value is much longer at 37°C. The *rrp46-L191H* cells also show significantly increased doubling time at both temperatures. Consistent with the solid media assays, the other *rrp* mutant variants modeling the pathogenic amino acid substitutions do not exhibit significantly different doubling times compared to the *BY4741* cells. Images shown are from a singular solid media growth assay with all samples plated on the same -Leu media plate. Data is representative of three independent experiments (n = 3).

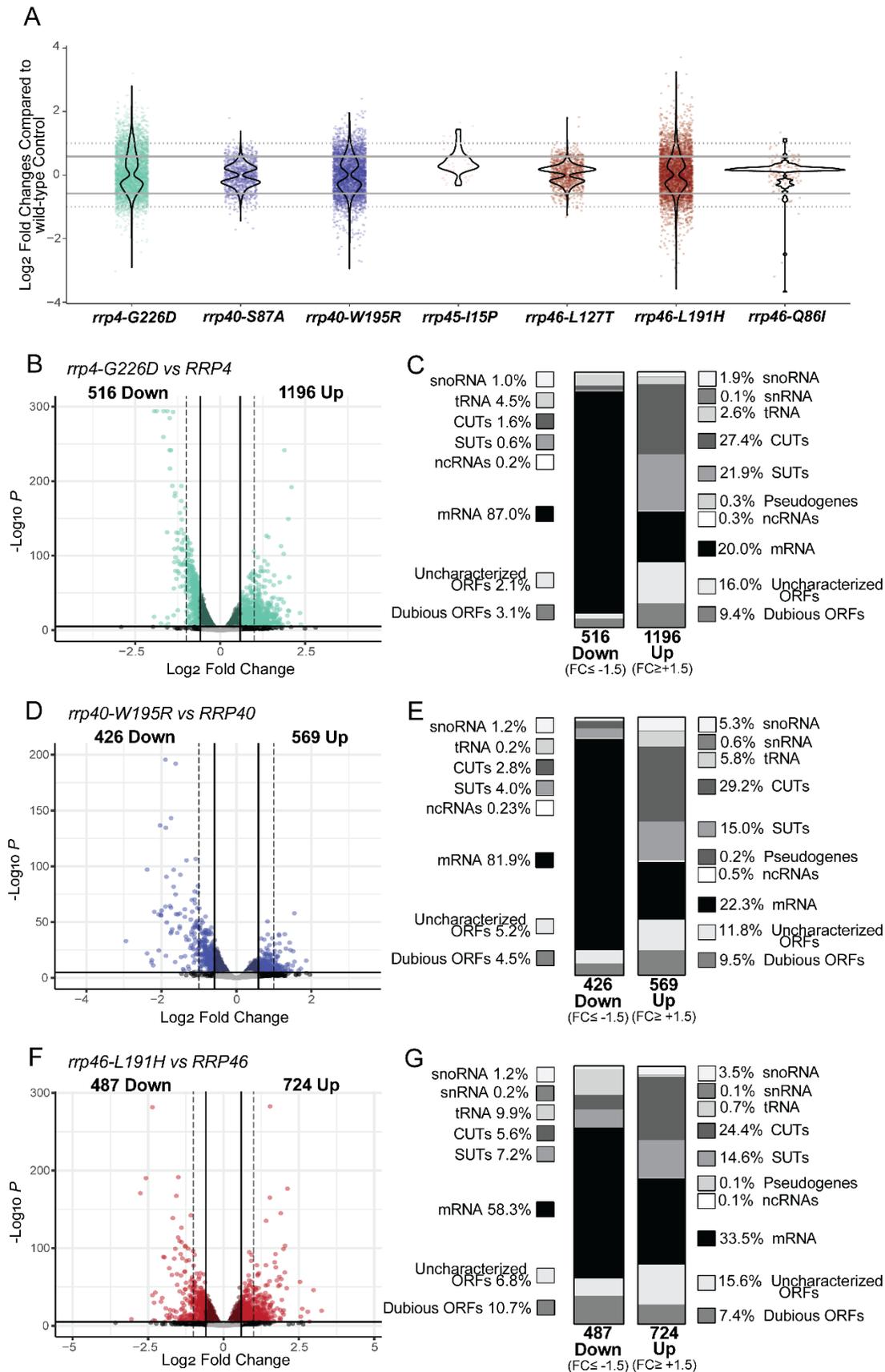


Figure 3. RNA-Seq analysis of *rrp* RNA exosomopathy mutant models reveal distinct transcriptomic changes in the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells.

RNA-Seq data was collected and analyzed as described in *Materials and Methods*. (A) Violin plots showing the distribution of transcripts identified in differential analysis as significant ($p < 0.05$) in each *rrp* mutant compared to the corresponding *RRP* wild-type control. The y-axis are the Log₂ Fold Change (LFC) values for each transcript. The solid grey line demarcates a Fold Change of +1.5 or -1.5 (LFC=0.5849625 or -0.5849625). The dotted grey line demarcates a Fold Change of +2 or -2 (LFC=1 or -1). These distributions reveal that the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells have the most transcripts with Fold Change $\geq +1.5$ (LFC ≥ 0.5849625) or ≤ -1.5 (LFC ≤ -0.5849625). (B-F) Volcano plots of the differentially expressed transcripts and classification of RNA types in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells. Classification of RNA types were calculated as percentages and include messenger RNA (mRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), transfer RNA (tRNA), cryptic unstable transcripts (CUTs; small, non-coding RNA), stable unannotated transcripts (SUTs; small, non-coding RNA), other non-coding RNA (ncRNA; e.g. *TLCl*), pseudogenes, and uncharacterized or dubious open reading frames (ORFs). (B) Differential expression analysis of *rrp4-G226D* cells compared to *RRP4* cells reveal 516 transcripts are significantly ($p < 0.05$) decreased (Down) and 1196 transcripts are significantly increased (Up) by 1.5-fold or more in the *rrp* mutant cells. Vertical lines demarcate FC values of ± 1.5 (straight line) and ± 2 (dotted line). (C) Stacked bar of the percentages of different RNA classes within the 516 Down and 1196 Up transcripts in *rrp4-G226D* cells reveal that decreased transcripts are predominantly mRNAs while the largest percentage of increased transcripts are ncRNAs CUTs and SUTs. (D) Differential expression analysis of *rrp40-W195R* cells compared to *RRP40* cells reveals 426 transcripts are significantly ($p < 0.05$) decreased (Down) and 569 transcripts are significantly increased (Up) by 1.5-fold or more within the *rrp* mutant cells. Vertical lines demarcate FC values of ± 1.5 (straight line) and ± 2 (dotted line). (E) Stacked bar of the percentages of different RNA classes within the 426 Down and 569 Up transcripts in *rrp40-W195R* cells reveal that decreased transcripts are predominantly mRNAs while increased transcripts

are majority ncRNAs. Within the majority ncRNAs decreased in *rrp40-W195R* cells, CUTs are nearly double the percentage as SUTs. (F) Differential expression analysis of *rrp46-L191H* cells compared to *RRP46* cells reveals 487 transcripts that are significantly ($p < 0.05$) decreased (Down) and 724 transcripts are significantly increased (Up) by 1.5-fold or more in the *rrp* mutant cells. Vertical lines demarcate FC values of ± 1.5 (straight line) and ± 2 (dotted line). (G) Stacked bar of the percentages of different RNA classes within the 487 Down transcripts in *rrp46-L191H* cells reveal that majority mRNAs, though to a lesser amount than in the other mutant models *rrp4-G226D* and *rrp40-W195R*. There is also a large percentage of decreased transcripts that classify as tRNAs within the *rrp46-L191H* cells. Stacked bar of the percentages of different RNA classes within the 724 Up transcripts in *rrp46-L191H* cells reveal a plurality between coding transcripts and ncRNAs, with about 1/3 of transcripts classifying as mRNAs and another 1/3 combined classifying as CUTs or SUTs.

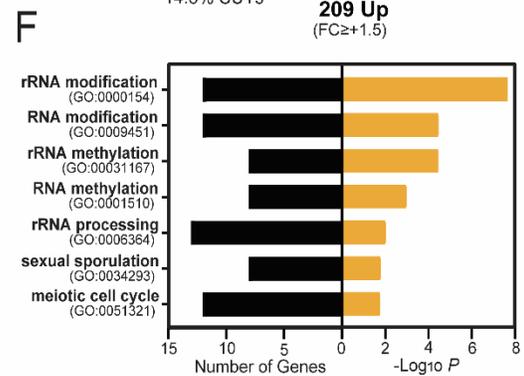
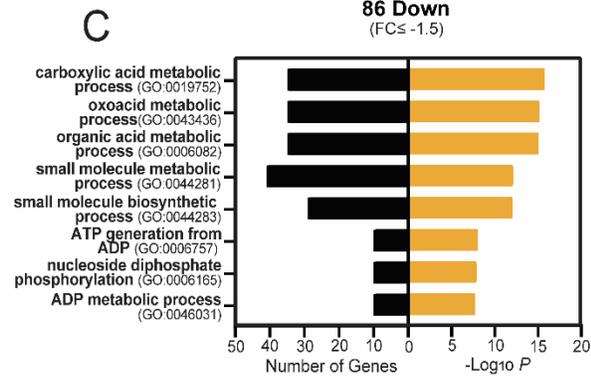
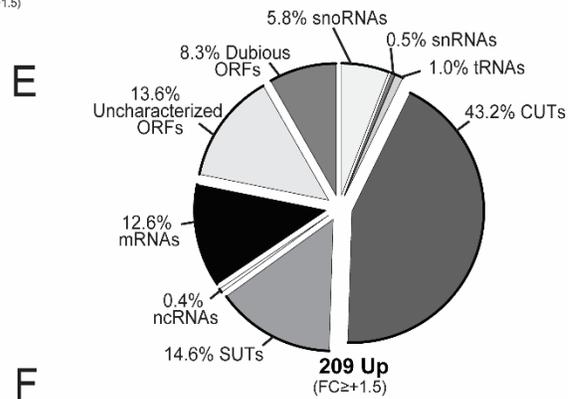
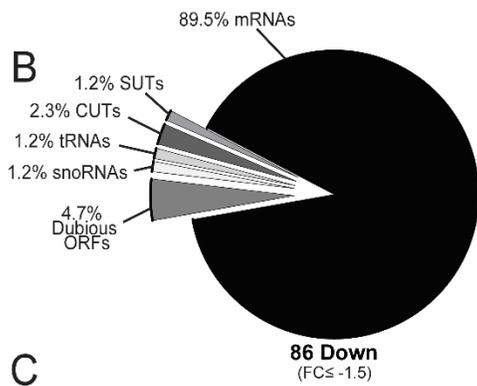
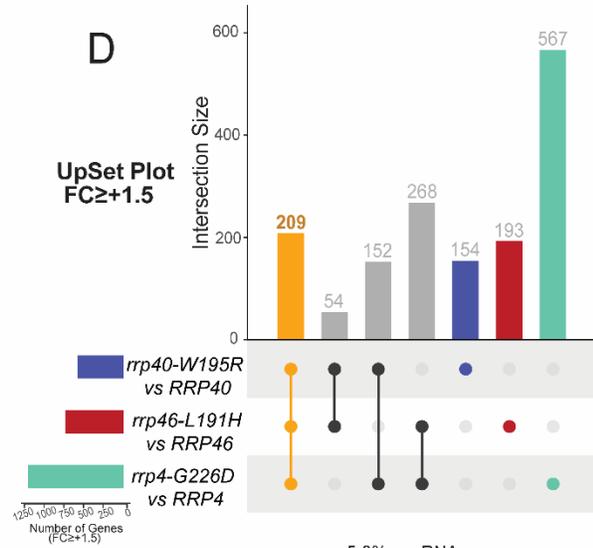
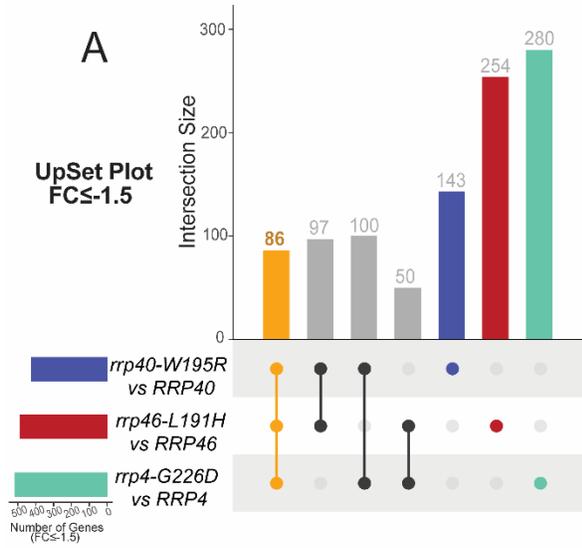


Figure 4. UpSet Plots of differentially expressed transcripts in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells reveal shared targets involved in metabolism and rRNA processing.

UpSet plots were generated with transcripts identified through differential expression analysis in the *rrp* mutant cells as significantly decreased 1.5 fold or more ($FC \leq -1.5$) and significantly increased 1.5 fold or more ($FC \geq +1.5$). UpSet intersections are shown as a matrix, with rows corresponding to the sets of samples (i.e. transcripts identified $FC \leq -1.5$ or $FC \geq +1.5$ within the three *rrp* mutant models) and columns corresponding to the intersection between these sets. The sets of samples are color coded; transcripts identified $FC \leq -1.5$ or $FC \geq +1.5$ from differential expression analysis of *rrp40-W195R* vs *RRP40* wild-type control cells are colored blue; transcripts identified $FC \leq -1.5$ or $FC \geq +1.5$ from differential expression analysis of *rrp46-L191H* vs *RRP46* wild-type control cells are colored red; transcripts identified $FC \leq -1.5$ or $FC \geq +1.5$ from differential expression analysis of *rrp4-G226D* vs *RRP4* wild-type control cells are colored teal. (A) The UpSet plot of significantly decreased 1.5 fold or more ($FC \leq -1.5$) transcripts reveals 86 are co-occurring within the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* datasets (intersection colored orange). (B) Pie chart of the types of RNA that comprise the intersection of shared decreased transcripts (Down) show that a majority are classified as mRNAs. (C) Gene ontology (GO) analysis for biological process of these shared 86 decreased transcripts reveal that metabolic and biosynthetic processes and are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated *p*-value for each GO term. (D) The UpSet plot of significantly increased 1.5 fold or more ($FC \geq +1.5$) transcripts reveals 209 are co-occurring within the *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* datasets (intersection colored orange). (E) Pie chart of the types of RNA that comprise the intersection of shared increased transcripts (Up) show that a majority are classified as CUTs or SUTs. (F) Gene ontology (GO) analysis for biological process of these shared 209 increased transcripts reveal that rRNA related processes and are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent

the $-\log$ of the associated p -value for each GO term. All GO analyses was performed the YeastMine web server on coding (mRNA) and non-coding RNA (tRNAs, snoRNAs, and snRNAs) classes.

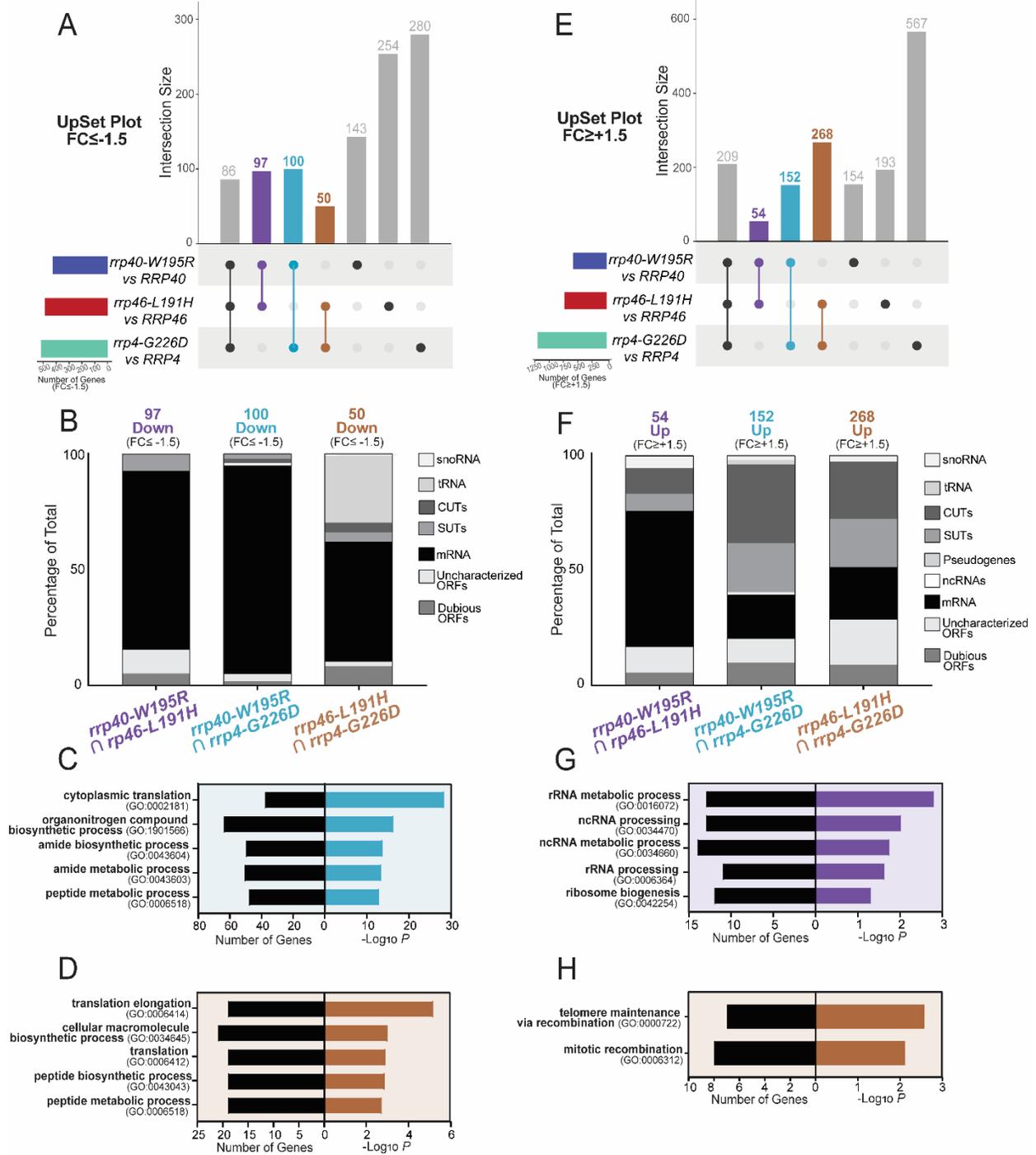


Figure 5. UpSet Plots of differentially expressed transcripts in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells reveal targets shared differently between dual combinations of the three *rrp* mutants.

UpSet plots generated as described in Figure 3. The intersections assessed here are transcripts either significantly decreased or increased 1.5 fold or more in both *rrp40-W195R* and *rrp46-L191H* (purple), *rrp40-W195R* and *rrp4-G226D* (light blue), or *rrp4-G226D* and *rrp46-L191H* (brown). (A) The UpSet plot of significantly decreased 1.5 fold or more ($FC \leq -1.5$) transcripts reveals 97 co-occurring within both *rrp40-W195R* and *rrp46-L191H* datasets, 100 co-occurring within both *rrp40-W195R* and *rrp4-G226D* datasets, and 50 co-occurring within both *rrp4-G226D* and *rrp46-L191H* datasets. (B) Stacked bar percentages of the RNA types that comprise the intersections of decreased transcripts (Down) identified between the *rrp40-W195R* and *rrp46-L191H* datasets, the *rrp40-W195R* and *rrp4-G226D* datasets, and the *rrp4-G226D* and *rrp46-L191H* datasets. (C) Gene ontology (GO) analysis for biological process of the 100 decreased transcripts co-occurring in the *rrp40-W195R* and *rrp4-G226D* datasets reveal that cellular translation and biosynthetic processes and are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (D) GO analysis for biological process of the 50 decreased transcripts co-occurring in the *rrp4-G226D* and *rrp46-L191H* datasets reveal that translation elongation and biosynthetic processes and are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (E) A) The UpSet plot of significantly increased 1.5 fold or more ($FC \geq +1.5$) transcripts reveals reveals 54 co-occurring within both *rrp40-W195R* and *rrp46-L191H* datasets, 162 co-occurring within both *rrp40-W195R* and *rrp4-G226D* datasets, and 268 co-occurring within both *rrp4-G226D* and *rrp46-L191H* datasets. (F) Stacked bar percentages of the RNA types that comprise the intersections of increased transcripts (Up) identified between the *rrp40-W195R* and *rrp46-L191H* datasets, the *rrp40-W195R* and *rrp4-G226D* datasets, and the *rrp4-G226D* and *rrp46-L191H* datasets. (G) GO analysis for biological process of the 54 increased

transcripts co-occurring in the *rrp40-W195R* and *rrp46-L191H* datasets reveal that processes involving rRNA and ncRNA processing and ribosome biogenesis are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (H) GO analysis for biological process of the 268 increased transcripts co-occurring in the *rrp4-G226D* and *rrp46-L191H* datasets reveal that telomere maintenance and mitotic processes and are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. All GO analyses was performed the YeastMine web server on coding (mRNA) and non-coding RNA (tRNAs, snoRNAs, and snRNAs) classes.

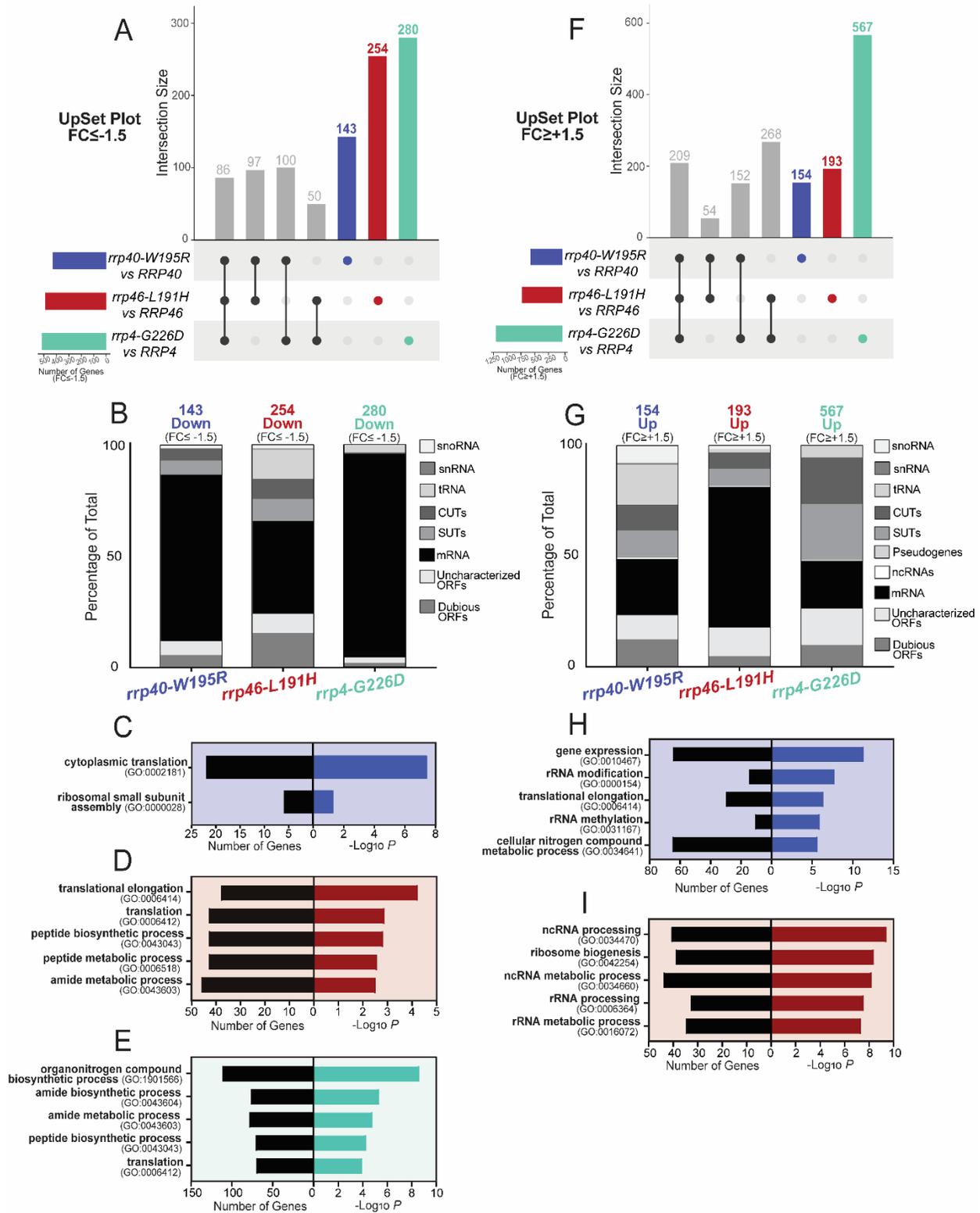


Figure 6. UpSet Plots of differentially expressed transcripts in *rrp4-G226D*, *rrp40-W195R* and *rrp46-L191H* cells reveal targets uniquely impacted in each of the three *rrp* mutants.

UpSet plots generated as described in Figure 3. The intersections assessed here are transcripts significantly decreased or increased 1.5 fold or more only in the *rrp40-W195R* dataset (blue), the *rrp46-L191H* dataset (red) or the *rrp4-G226D* dataset (teal). (A) The UpSet plot of significantly decreased 1.5 fold or more ($FC \leq -1.5$) transcripts reveals 143 occurring solely in the *rrp40-W195R* dataset, 100 occurring solely in the *rrp46-L191H* dataset, and 280 occurring solely in the *rrp4-G226D* dataset. (B) Stacked bar percentages of the RNA types that comprise the decreased transcripts (Down) identified only within the *rrp40-W195R*, *rrp46-L191H* or *rrp4-G226D* datasets. (C) Gene ontology (GO) analysis for biological process of the 143 decreased transcripts occurring only in the *rrp40-W195R* dataset reveals that cytoplasmic translation and ribosomal small unit assembly processes are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (D) GO analysis for biological process of the 254 decreased transcripts occurring only in the *rrp46-L191H* dataset reveals processes related to translation and biomolecular synthesis are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (E) GO analysis for biological process of the 280 decreased transcripts occurring only in the *rrp4-G226D* dataset reveals processes related to biosynthesis are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (F) The UpSet plot of significantly increased 1.5 fold or more ($FC \geq +1.5$) transcripts reveals 154 occurring solely in the *rrp40-W195R* dataset, 193 occurring solely in the *rrp46-L191H* dataset, and 567 occurring solely in the *rrp4-G226D* dataset. (G) Stacked bar percentages of the RNA types that comprise the increased transcripts (Up) identified only within the *rrp40-W195R*, *rrp46-L191H* or *rrp4-G226D* datasets. (H) GO analysis for biological process of the 154 increased transcripts occurring only in the *rrp40-W195R* dataset reveals that processes related to gene expression, rRNA

modifications and translation are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. (I) GO analysis for biological process of the 193 increased transcripts occurring only in the *rrp46-L191H* dataset reveals processes related to rRNA processing and ribosome biogenesis are significantly enriched. Black bars represent the number of transcripts that are linked to each biological process category. Orange bars represent the $-\log$ of the associated p -value for each GO term. All GO analyses was performed the YeastMine web server on coding (mRNA) and non-coding RNA (tRNAs, snoRNAs, and snRNAs) classes.

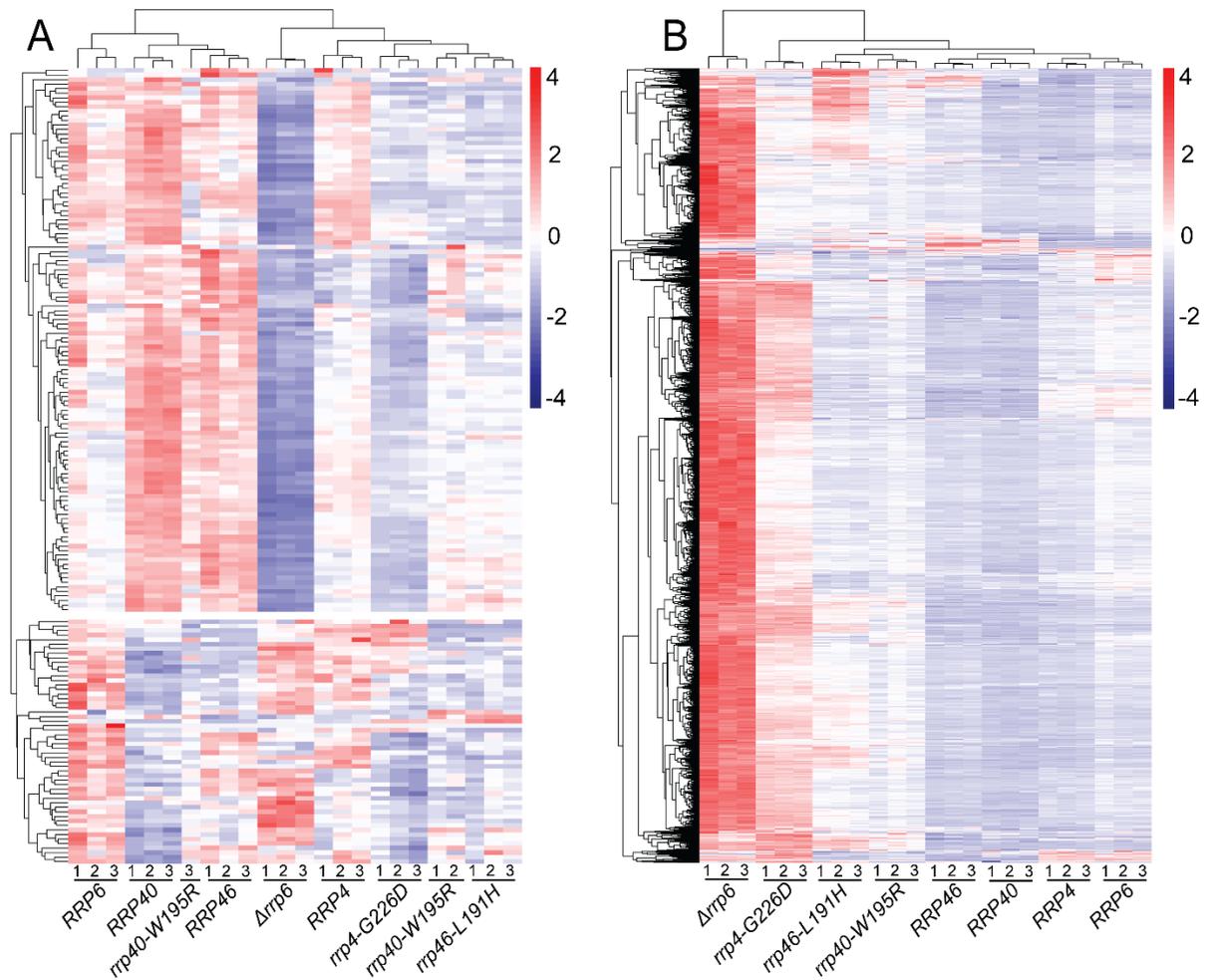


Figure 7. Heatmaps of *rrp* mutants reveal broad changes in ribosomal protein gene and CUTs/SUTs expression.

Heatmaps were generated on with FPKM estimates of (A) all annotated ribosomal protein *RPS* and *RPL* genes and (B) all annotated CUTs and SUTs [70]. FPKM estimates were calculated as described in *Materials and Methods*. Gene expression estimates are scaled for heatmap visualization and coloring is a gradient of higher (red) to lower (blue) scaled values.

4.7 Chapter IV Supplementary Materials

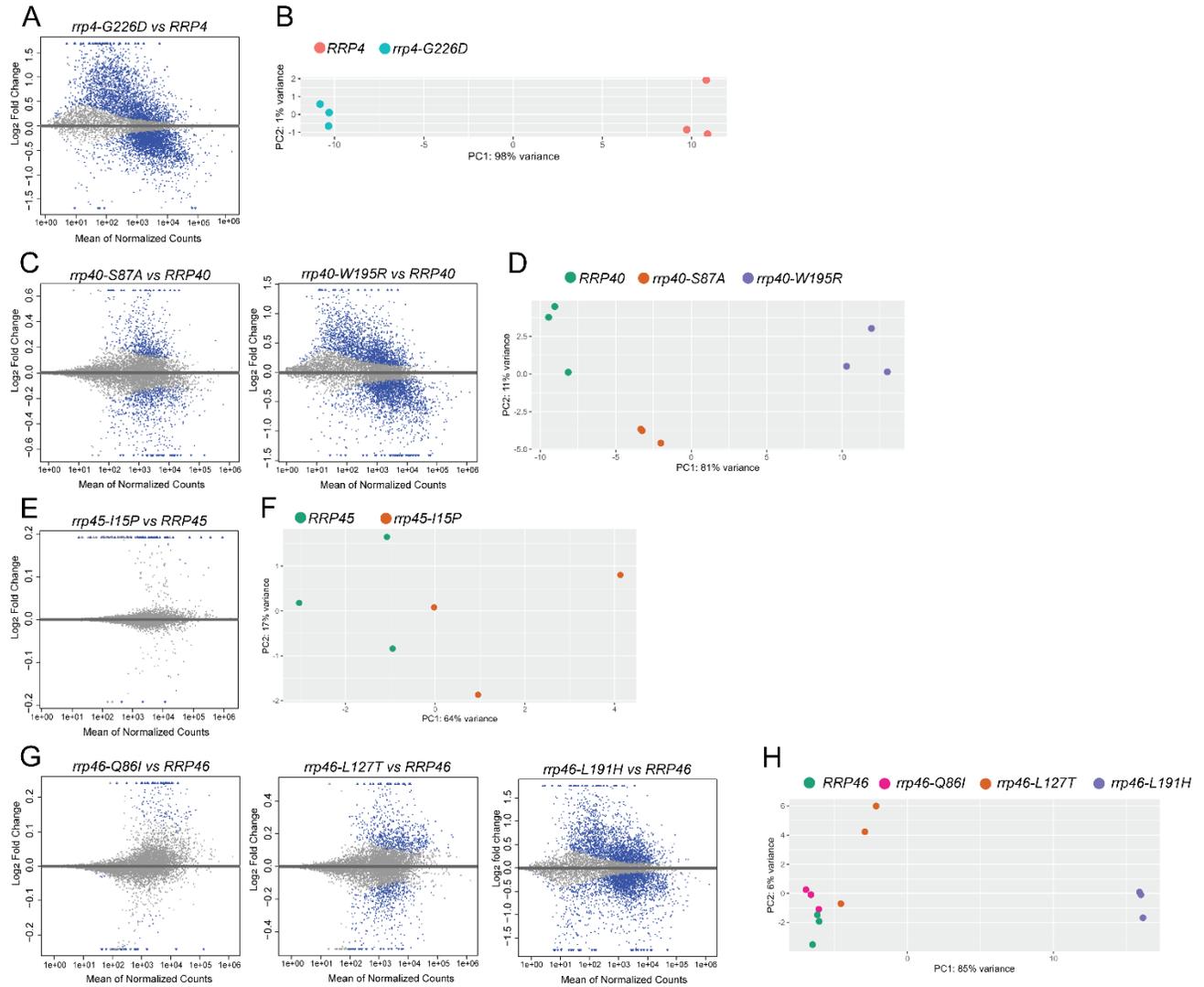
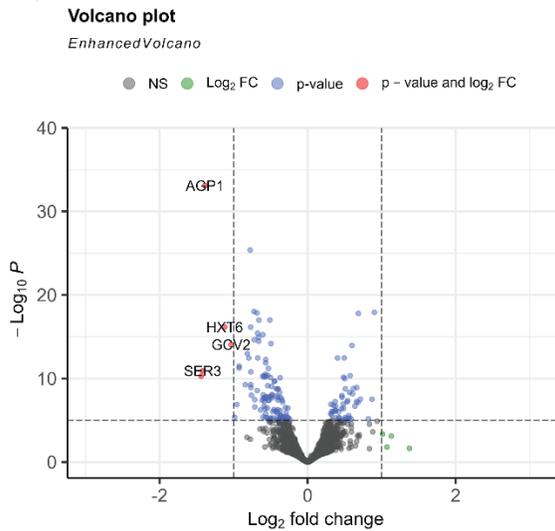


Figure S1. MA plots for differential expression analysis results and PCA plots of sample clustering in RNA-Seq experiment.

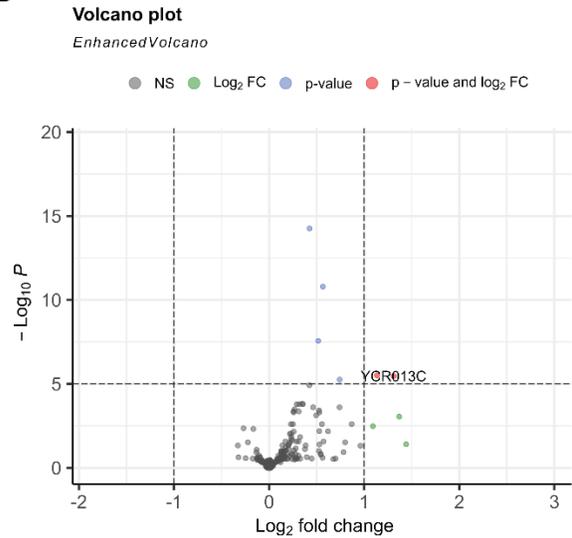
MA plots were generated on DESeq2 results and plot the log₂ Fold Change (FC) values attributable to each transcript over the mean of normalized gene counts. Significance ($p < 0.05$) is indicated by blue coloring of data points. PCA Principal component analysis (PCA) of RNA-Seq data collected from triplicate *rrp* or *RRP* samples as indicated. (A) MA plots of the DESeq2 results comparing *rrp4-G226D* to wild-type control

RRP4 samples. (B) PCA analysis reveals clustering of the *rrp4-G226D* samples away from the *RRP4* samples, suggesting that gene expression patterns of the two genotypes are distinct. (C) MA plots of the DESeq2 results comparing *rrp40* mutants to wild-type control *RRP40* samples. (D) PCA analysis reveals clustering of the *rrp40* mutant samples away from the *RRP40* samples, suggesting that gene expression patterns of the genotypes are distinct. The *rrp40-W195R* sample however clusters the furthest from the *RRP40* samples, encompassing 81% variance. (E) MA plots of the DESeq2 results comparing *rrp45-I15P* to wild-type control *RRP4* samples. (F) PCA analysis reveals that the *rrp45-I15P* mutant samples do not independently cluster away from the *RRP45* samples, suggesting that gene expression patterns of these genotypes are not distinct. (G) MA plots of the DESeq2 results comparing *rrp46* mutants to wild-type control *RRP46* samples. (H) PCA analysis reveals clustering of the *rrp46-L191H* mutant samples away from the *RRP40* samples, suggesting that gene expression patterns of those genotypes are distinct. The *rrp46-Q86I* and *rrp46-L127T* samples do not cluster independently from the *RRP46* samples, suggesting the gene expression patterns of these *rrp46* mutants are very similar to that of the wild-type control genotype.

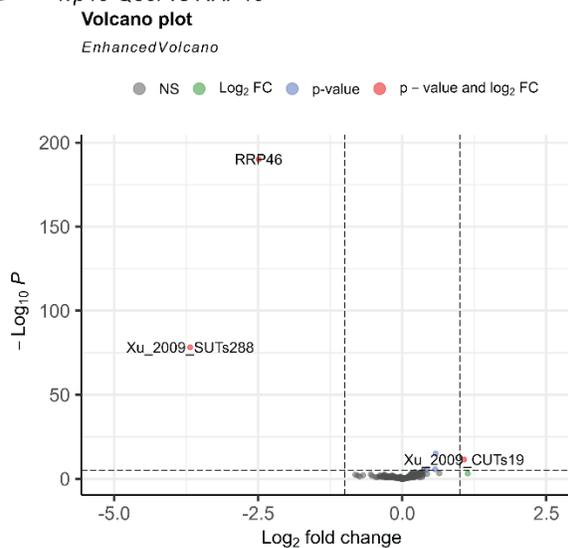
A *rrp40-S87A vs RRP40*



B *rrp45-I15P vs RRP45*



C *rrp46-Q86I vs RRP46*



D *rrp46-L127T vs RRP46*

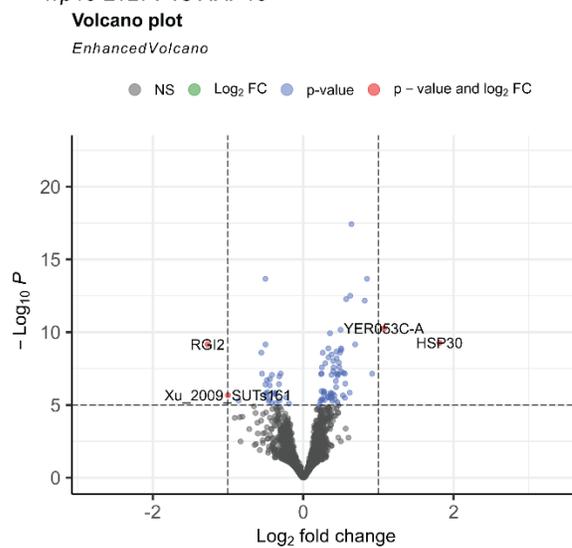


Figure S2. Volcano plots of differentially expressed transcripts identified in *rrp40-S87A*, *rrp45-I15P*, *rrp46-Q86I* and *rrp46-L127T* samples.

Differential expression analysis was performed for each *rrp* mutant sample compared to its corresponding wild-type control *RRP* sample. Volcano plots were generated as described in *Materials and Methods*. Coloring is consistent across each volcano plot; grey represents transcripts that are not significantly different between the *rrp* and *RRP* samples; green represents transcripts that are identified as increased or

decreased by 1.5 fold change or more in the *rrp* sample compared to the *RRP* sample; blue represents transcripts that are identified as significantly different ($p < 0.05$) in the *rrp* sample compared to the *RRP* sample; red represents transcripts that are identified as both significantly different ($p < 0.05$) in the *rrp* sample compared to the *RRP* sample and increased or decreased by 1.5 fold change or more. (A) Volcano plot of differentially expressed transcripts in *rrp40-S87A* compared to *RRP40* wild-type control reveal several that are statistically significant but few that are increased or decreased by 1.5 fold change or more. (B) Volcano plot of differentially expressed transcripts in *rrp45-I15P* compared to *RRP45* reveal very few differences between the *rrp* mutant and wild-type control. (C) Volcano plot of differentially expressed transcripts in *rrp46-Q86I* compared to *RRP46* reveal essentially no difference between the *rrp* mutant and wild-type control. (D) Volcano plot of differentially expressed transcripts in *rrp46-L127T* compared to *RRP46* reveal several that are statistically significant but few that are increased or decreased by 1.5 fold change or more.

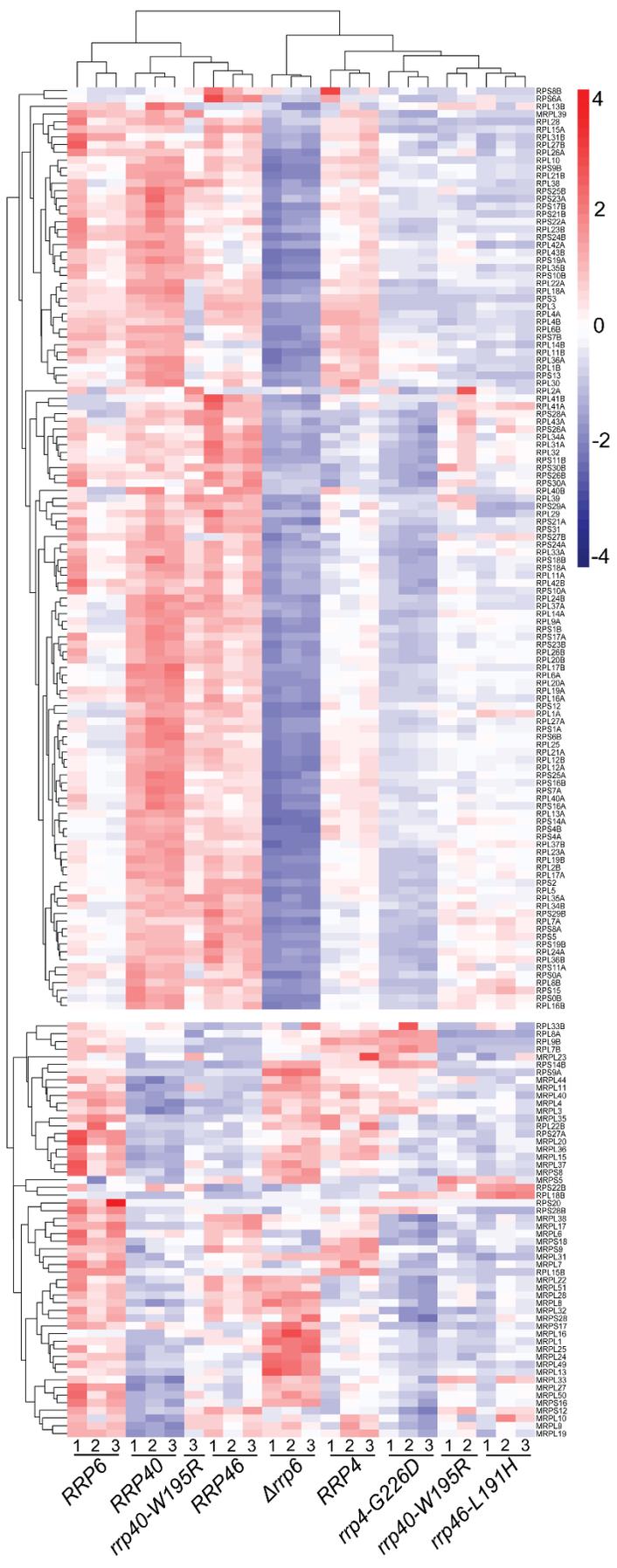


Figure S3. Heatmap of ribosomal protein genes with gene names.

Heatmaps were generated on with FPKM estimates of all annotated ribosomal protein *RPS* and *RPL* genes. FPKM estimates were calculated as described in *Materials and Methods*. Gene expression estimates are scaled for heatmap visualization and coloring is a gradient of higher (red) to lower (blue) scaled values. The gene names of each *RPS* or *RPL* transcript are listed on the right.

Table S1. List of RNA exosomopathy mutations and associated pathologies

Gene (Human/ Budding Yeast)	Position in RNA exosome	Missense Mutation in Human	Model in Budding Yeast	Disease Pathology	Ref(s)
<i>EXOSC1/CSL4</i>	Cap subunit	S35L (homozygous)	A58L	PCH1F (OMIM #619304)	[48]
<i>EXOSC2/RRP4</i>	Cap subunit	G30V (homozygous); G198D, G30V (compound het);	G58V; G226D;	SHRF syndrome (OMIM #617763)	[113]
<i>EXOSC3/RRP40</i>	Cap subunit	G31A (homozygous); D132A (homozygous); W238R, G31A (compound het)	G8A; S87A; W195R	PCH1b (OMIM #614678)	[40, 44-46, 49, 110]
<i>EXOSC4/RRP41*</i>	Core subunit	L187P (homozygous)	L187P	Cerebellar atrophy	Collaboration with Dr. Almundher Al-Maawali
<i>EXOSC5/RRP46</i>	Core subunit	T114I, <i>EXOSC5</i> del (compound het); M148T (homozygous); L206H (homozygous)	Q86I; L127T; L191H	Cerebellar hypoplasia (OMIM #619576)	[47]
<i>EXOSC7/RRP42*</i>	Core subunit	S229L (homozygous); S229L,?; (Compound het)	S214L	PCH or cerebellar ataxia	Collaboration with Dr. Guoliang Chai and Dr. Lan Yu
<i>EXOSC8/RRP43</i>	Core subunit	A2V (homozygous); S272T (homozygous)	A2V; S389T	PCH1c (OMIM #616081)	[41]
<i>EXOSC9/RRP45</i>	Core subunit	L14P (homozygous); G51R (homozygous)	I15P; G52R	PCH1d (OMIM #618065)	[42, 50, 51]

*Entries with asterisks are from unpublished communication with listed collaborators

Table S2. *Saccharomyces cerevisiae* strains and plasmids

Strain	Description	Ref.
<i>rrp4Δ</i> (yAV1104)	<i>MATα his3Δ1 leu2Δ0 ura3Δ0 lys2Δ0 rrp4Δ::NEO [RRP4,URA3]</i>	[62]
<i>rrp40Δ</i> (yAV1107)	<i>MATα his3Δ1 leu2Δ0 ura3Δ0 rrp40Δ::NEO [RRP40,URA3]</i>	[62]
<i>rrp45Δ</i> (yAV1410)	<i>MATα his3Δ1 leu2Δ0 ura3Δ0 rrp45Δ::NEO [RRP40,URA3]</i>	[62]
<i>rrp46Δ</i> (yAV1105)	<i>MATα his3Δ1 leu2Δ0 ura3Δ0 lys2Δ0 rrp46Δ::KANMX [RRP4,URA3]</i>	[62]
BY4741 (ACY1051)	<i>MATα; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0</i>	Hansen
<i>rrp6Δ</i> (ACY1641)	<i>MATα; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; RRP6::kanMX4</i>	This Study
pRS315	<i>CEN6, LEU2, amp^R</i>	[63]
pAC3656	<i>RRP4-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[54]
pAC3659	<i>rrp4-G226D-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[54]
pAC3652	<i>RRP40-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[54]
pAC3654	<i>rrp40-S87A-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	This Study
pAC3655	<i>rrp40-W195R-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[54]
pAC3479	<i>RRP45-Native 3'UTR CEN6, LEU2, amp^R; pAV975</i>	This Study
pAC3480	<i>rrp45-I15P-Native 3'UTR, CEN6, LEU2, amp^R; pAV975</i>	This Study
pAC3482	<i>RRP46-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[47]
pAC3483	<i>rrp46-Q86I-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[47]
pAC3484	<i>rrp46-L191H-Native 3'UTR in pRS315, CEN6, LEU2, amp^R</i>	[47]

pAC3534	<i>rrp46-L127T-Native 3'UTR</i> in pRS315, <i>CEN6</i> , <i>LEU2</i> , <i>amp^R</i>	[47]
pAC3752	<i>RRP6</i> in pRS315, <i>CEN6</i> , <i>LEU2</i> , <i>amp^R</i>	This Study

Supplemental Documentation S1. Full RNA-Seq Datasets

https://emory-my.sharepoint.com/:x:/g/personal/msterre_emory_edu/EZG1VZuvbq1OodKJzqOcuWABJGFIRWyKegbPG7iPdGAYkA?e=ZIWBNP

Supplemental Documentation S2. Full lists of differentially expressed genes ($\geq+1.5$ or ≤-1.5 Fold Change [FC], $p<0.05$)

https://emory-my.sharepoint.com/:x:/g/personal/msterre_emory_edu/EZG1VZuvbq1OodKJzqOcuWABJGFIRWyKegbPG7iPdGAYkA?e=ZIWBNP

Supplemental Documentation S3. Full Gene Ontology (GO) Analyses Terms Lists

GO Enrichment Biological process *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* DOWN targets

carboxylic acid metabolic process 1.9132938057834414e-16
YBR026C, YBR145W, YBR196C, YCL030C, YCR012W, YDR158W, YDR516C, YER055C, YER069W, YER073W, YER091C, YFR047C, YFR053C, YGL253W, YGR192C, YGR286C, YHR018C, YHR174W, YHR208W, YJR009C, YJR016C, YKL060C, YKL152C, YLR058C, YLR355C, YMR083W, YMR116C, YNL040W, YNL104C, YNR058W, YOL086C, YOL126C, YOR184W, YOR375C, YPL061W GO:0019752

oxoacid metabolic process 7.7278883114513e-16
YBR026C, YBR145W, YBR196C, YCL030C, YCR012W, YDR158W, YDR516C, YER055C, YER069W, YER073W, YER091C, YFR047C, YFR053C, YGL253W, YGR192C, YGR286C, YHR018C, YHR174W, YHR208W, YJR009C, YJR016C, YKL060C, YKL152C, YLR058C, YLR355C, YMR083W, YMR116C, YNL040W, YNL104C, YNR058W, YOL086C, YOL126C, YOR184W, YOR375C, YPL061W GO:0043436

organic acid metabolic process 8.985756181260476e-16
YBR026C, YBR145W, YBR196C, YCL030C, YCR012W, YDR158W, YDR516C, YER055C, YER069W, YER073W, YER091C, YFR047C, YFR053C, YGL253W, YGR192C, YGR286C, YHR018C

,YHR174W,YHR208W,YJR009C,YJR016C,YKL060C,YKL152C,YLR058C,YLR355C,YMR083W,YMR116C,YNL040W,YNL104C,YNR058W,YOL086C,YOL126C,YOR184W,YOR375C,YPL061W GO:0006082

small molecule metabolic process 8.483622251500631e-13

YBR026C,YBR145W,YBR196C,YCL030C,YCR012W,YDL055C,YDR158W,YDR516C,YER055C,YER069W,YER073W,YER091C,YFR047C,YFR053C,YGL253W,YGR192C,YGR286C,YHR018C,YHR174W,YHR208W,YJR009C,YJR016C,YKL060C,YKL152C,YKL216W,YLR058C,YLR355C,YLR359W,YLR420W,YMR083W,YMR116C,YMR120C,YNL040W,YNL104C,YNR058W,YOL086C,YOL126C,YOL143C,YOR184W,YOR375C,YPL061W GO:0044281

small molecule biosynthetic process 9.445063099779353e-13

YBR026C,YBR145W,YBR196C,YCL030C,YCR012W,YDR158W,YER055C,YER069W,YER073W,YER091C,YGR192C,YGR286C,YHR018C,YHR208W,YJR009C,YJR016C,YKL060C,YKL152C,YLR058C,YLR355C,YMR083W,YNL104C,YNR058W,YOL086C,YOL126C,YOL143C,YOR184W,YOR375C,YPL061W GO:0044283

glycolytic process 9.56019356913603e-9

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0006096

ATP generation from ADP 9.56019356913603e-9

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0006757

nucleoside diphosphate phosphorylation 1.4179575247086696e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0006165

nucleotide phosphorylation 1.4179575247086696e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0046939

ADP metabolic process 2.069714845730972e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0046031

purine nucleoside diphosphate metabolic process 2.9773540332607258e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0009135

purine ribonucleoside diphosphate metabolic process

2.9773540332607258e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0009179

ribonucleoside diphosphate metabolic process 5.91927626797057e-8

YBR196C,YCR012W,YDR516C,YFR053C,YGL253W,YGR192C,YHR174W,YJR009C,YKL060C,YKL152C GO:0009185

nucleoside diphosphate metabolic process 2.0427668562908302e-7
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C GO:0009132

organic acid biosynthetic process 3.4696678547261545e-7
YBR026C, YCL030C, YDR158W, YER055C, YER069W, YER073W, YER091C, YGR286C, Y
HR018C, YHR208W, YJR016C, YLR058C, YLR355C, YNL104C, YNR058W, YOR184W, YOR375C
, YPL061W GO:0016053

carboxylic acid biosynthetic process 3.4696678547261545e-7
YBR026C, YCL030C, YDR158W, YER055C, YER069W, YER073W, YER091C, YGR286C, Y
HR018C, YHR208W, YJR016C, YLR058C, YLR355C, YNL104C, YNR058W, YOR184W, YOR375C
, YPL061W GO:0046394

pyruvate metabolic process 3.581924878723665e-7
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C GO:0006090

amino acid metabolic process 0.000002028996865567142
YBR145W, YCL030C, YDR158W, YER055C, YER069W, YER091C, YFR047C, YHR018C, Y
HR208W, YJR016C, YLR058C, YLR355C, YMR083W, YMR116C, YNL040W, YNL104C, YOL086C
, YOR184W, YOR375C GO:0006520

ribonucleotide metabolic process 0.000004492238135530351
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C, YKL216W, YLR359W, YLR420W, YMR120C GO:0009259

ATP metabolic process 0.00000478767674557079
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C GO:0046034

glucose metabolic process 0.000007147385827293535
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YJR009C, YKL060C, Y
KL152C, YOL126C GO:0006006

ribose phosphate metabolic process 0.00000925501081569874
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C, YKL216W, YLR359W, YLR420W, YMR120C GO:0019693

nucleobase-containing small molecule metabolic process
0.000011028234466871715
YBR196C, YCR012W, YDL055C, YDR516C, YFR047C, YFR053C, YGL253W, YGR192C, Y
HR174W, YJR009C, YKL060C, YKL152C, YKL216W, YLR359W, YLR420W, YMR120C, YOR184W
, YPL061W GO:0055086

purine ribonucleoside triphosphate metabolic process
0.000012624823784690073
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, Y
KL060C, YKL152C GO:0009205

generation of precursor metabolites and energy
0.000016259430053098087
YBR026C, YBR145W, YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YNL031C, YOL086C, YOL126C, YPL061W, YPR160W
GO:0006091

purine nucleoside triphosphate metabolic process
0.000018095002197382425
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C GO:0009144

monocarboxylic acid metabolic process 0.000021197276826081024
YBR026C, YBR196C, YCR012W, YDR516C, YER073W, YFR053C, YGL253W, YGR192C, YGR286C, YHR174W, YJR009C, YKL060C, YKL152C, YNR058W, YPL061W GO:0032787

purine-containing compound metabolic process
0.000031904043384424937
YBR196C, YCR012W, YDR516C, YFR047C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YLR359W, YMR120C, YOR184W, YPL061W GO:0072521

ribonucleoside triphosphate metabolic process
0.000035675311368569545
YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C GO:0009199

nucleotide metabolic process 0.000043675426667829436
YBR196C, YCR012W, YDR516C, YFR047C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YKL216W, YLR359W, YLR420W, YMR120C, YPL061W
GO:0009117

amino acid biosynthetic process 0.00007019179403866245
YCL030C, YDR158W, YER055C, YER069W, YER091C, YHR018C, YHR208W, YJR016C, YLR058C, YLR355C, YNL104C, YOR184W, YOR375C GO:0008652

organic substance biosynthetic process 0.00007444026314192981
YBR026C, YBR031W, YBR145W, YBR196C, YCL030C, YCR012W, YDL055C, YDL061C, YDR012W, YDR064W, YDR158W, YER055C, YER069W, YER073W, YER091C, YFR047C, YGL030W, YGL123W, YGL253W, YGR148C, YGR192C, YGR286C, YHL015W, YHR018C, YHR208W, YIL119C, YJR009C, YJR016C, YJR047C, YKL060C, YKL152C, YKL216W, YLR029C, YLR058C, YLR167W, YLR355C, YLR359W, YLR420W, YMR083W, YMR116C, YMR120C, YMR136W, YNCK0005C, YNL031C, YNL040W, YNL098C, YNL104C, YNL178W, YNR058W, YOL086C, YOL126C, YOL143C, YOR063W, YOR184W, YOR375C, YPL061W GO:1901576

nucleoside phosphate metabolic process 0.00007601987111217217
YBR196C, YCR012W, YDR516C, YFR047C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YKL216W, YLR359W, YLR420W, YMR120C, YPL061W
GO:0006753

purine nucleotide metabolic process 0.00008854535566664348
 YBR196C, YCR012W, YDR516C, YFR047C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YLR359W, YMR120C, YPL061W GO:0006163

purine ribonucleotide metabolic process 0.00009938332392313004
 YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YLR359W, YMR120C GO:0009150

biosynthetic process 0.00009946589430705813
 YBR026C, YBR031W, YBR145W, YBR196C, YCL030C, YCR012W, YDL055C, YDL061C, YDR012W, YDR064W, YDR158W, YER055C, YER069W, YER073W, YER091C, YFR047C, YGL030W, YGL123W, YGL253W, YGR148C, YGR192C, YGR286C, YHL015W, YHR018C, YHR208W, YIL119C, YJR009C, YJR016C, YJR047C, YKL060C, YKL152C, YKL216W, YLR029C, YLR058C, YLR167W, YLR355C, YLR359W, YLR420W, YMR083W, YMR116C, YMR120C, YMR136W, YNCK0005C, YNL031C, YNL040W, YNL098C, YNL104C, YNL178W, YNR058W, YOL086C, YOL126C, YOL143C, YOR063W, YOR184W, YOR375C, YPL061W GO:0009058

nucleoside triphosphate metabolic process 0.00012044996163161632
 YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C GO:0009141

hexose metabolic process 0.00012044996163161632
 YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YJR009C, YKL060C, YKL152C, YOL126C GO:0019318

carbohydrate catabolic process 0.00012203603992794198
 YBR056W, YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YHR174W, YJR009C, YKL060C, YKL152C, YPR160W GO:0016052

monosaccharide metabolic process 0.0005056012136125759
 YBR196C, YCR012W, YDR516C, YFR053C, YGL253W, YGR192C, YJR009C, YKL060C, YKL152C, YOL126C GO:0005996

gluconeogenesis 0.0006469246313186214
 YBR196C, YCR012W, YGR192C, YJR009C, YKL060C, YKL152C, YOL126C GO:0006094

hexose biosynthetic process 0.0006469246313186214
 YBR196C, YCR012W, YGR192C, YJR009C, YKL060C, YKL152C, YOL126C GO:0019319

monosaccharide biosynthetic process 0.0010279344491411564
 YBR196C, YCR012W, YGR192C, YJR009C, YKL060C, YKL152C, YOL126C GO:0046364

alpha-amino acid biosynthetic process 0.0017427734848315631
 YDR158W, YER069W, YER091C, YHR018C, YHR208W, YJR016C, YLR058C, YLR355C, YNL104C, YOR184W, YOR375C GO:1901607

organonitrogen compound metabolic process 0.0019635638835045397
YBR031W, YBR145W, YBR158W, YBR196C, YCL030C, YCR012W, YDL055C, YDL061C, YDR012W, YDR064W, YDR158W, YDR516C, YER055C, YER069W, YER091C, YFR047C, YFR053C, YGL030W, YGL123W, YGL253W, YGR086C, YGR148C, YGR192C, YGR286C, YHL015W, YHR018C, YHR174W, YHR208W, YJR009C, YJR016C, YJR047C, YKL060C, YKL152C, YKL216W, YLR029C, YLR058C, YLR167W, YLR355C, YLR359W, YLR420W, YMR083W, YMR116C, YMR120C, YNCK0005C, YNL040W, YNL104C, YNL178W, YNR058W, YOL086C, YOL143C, YOR063W, YOR184W, YOR375C, YPL061W GO:1901564

organonitrogen compound biosynthetic process 0.003682952310260568
YBR031W, YCL030C, YDL055C, YDL061C, YDR012W, YDR064W, YDR158W, YER055C, YER069W, YER091C, YFR047C, YGL030W, YGL123W, YGR148C, YGR286C, YHL015W, YHR018C, YHR208W, YJR016C, YJR047C, YKL216W, YLR029C, YLR058C, YLR167W, YLR355C, YLR359W, YLR420W, YMR116C, YMR120C, YNCK0005C, YNL040W, YNL104C, YNL178W, YNR058W, YOL143C, YOR063W, YOR184W, YOR375C GO:1901566

branched-chain amino acid biosynthetic process 0.007751842444152609
YDR158W, YHR208W, YJR016C, YLR355C, YNL104C GO:0009082

fructose transmembrane transport 0.014882105203627196
YDR342C, YDR343C, YFR053C, YGL253W, YHR092C GO:0015755

alpha-amino acid metabolic process 0.01733312508640786
YDR158W, YER069W, YER091C, YFR047C, YHR018C, YHR208W, YJR016C, YLR058C, YLR355C, YNL104C, YOR184W, YOR375C GO:1901605

glucose transmembrane transport 0.04381636024535303
YDR342C, YDR343C, YFR053C, YGL253W, YHR092C GO:1904659

isoleucine biosynthetic process 0.04741397481032412
YDR158W, YHR208W, YJR016C, YLR355C GO:0009097

maintenance of translational fidelity 0.04741397481032412
YLR167W, YMR116C, YNL178W, YOR063W GO:1990145

GO Enrichment Biological process *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* UP targets

rRNA modification 3.13065138371405e-8
YNCB0003W, YNCC0011W, YNCD0010C, YNCE0001C, YNCE0003W, YNCG0013W, YNCG0026W, YNCH0014W, YNCN0014W, YNCN0018W, YNCO0006W, YNCO0015C GO:0000154

rRNA methylation 0.000045694204876823595
 YNCB0003W, YNCC0011W, YNCD0010C, YNCE0003W, YNCG0026W, YNCH0014W, YNCN0014W, YNCO0015C GO:0031167

RNA modification 0.00004948689175134433
 YNCB0003W, YNCC0011W, YNCD0010C, YNCE0001C, YNCE0003W, YNCG0013W, YNCG0026W, YNCH0014W, YNCN0014W, YNCN0018W, YNCO0006W, YNCO0015C GO:0009451

RNA methylation 0.0013578093074990004
 YNCB0003W, YNCC0011W, YNCD0010C, YNCE0003W, YNCG0026W, YNCH0014W, YNCN0014W, YNCO0015C GO:0001510

rRNA processing 0.013375679345983249
 YGR030C, YNCB0003W, YNCC0011W, YNCD0010C, YNCE0001C, YNCE0003W, YNCG0013W, YNCG0026W, YNCH0014W, YNCN0014W, YNCN0018W, YNCO0006W, YNCO0015C GO:0006364

sexual sporulation 0.02063120002983751
 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W GO:0034293

sexual sporulation resulting in formation of a cellular spore
 0.02063120002983751
 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W GO:0043935

macromolecule methylation 0.02325490328413725
 YNCB0003W, YNCC0011W, YNCD0010C, YNCE0003W, YNCG0026W, YNCH0014W, YNCN0014W, YNCO0015C GO:0043414

meiotic cell cycle 0.023922389921456625
 YBR180W, YDR403W, YDR523C, YHR157W, YLR445W, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W, YPL018W, YPL121C GO:0051321

sporulation resulting in formation of a cellular spore
 0.02733369108846444
 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W, YPL121C GO:0030435

sporulation 0.03300612656585292
 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W, YPL121C GO:0043934

anatomical structure formation involved in morphogenesis
 0.03789644415943166
 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W, YNL018C, YNL034W, YOR338W, YPL121C GO:0048646

GO Enrichment Biological process for *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* UP mRNA targets

meiotic cell cycle 0.010976593052401529
YBR180W, YDR403W, YDR523C, YHR157W, YLR445W, YMR017W, YMR306W, YPL018W, YPL121C GO:0051321

meiotic cell cycle process 0.0333326121184674
YBR180W, YDR403W, YDR523C, YHR157W, YLR445W, YMR017W, YMR306W, YPL121C
GO:1903046

cellular component morphogenesis 0.03513414786259377
YBR180W, YDR403W, YDR523C, YMR017W, YMR306W GO:0032989

cellular component assembly involved in morphogenesis
0.04607712625283756 YBR180W, YDR403W, YDR523C, YMR017W, YMR306W
GO:0010927

GO Enrichment Biological process *rrp4-G226D* alone DOWN targets

YKR059W, YLL018C, YLR069C, YLR150W, YLR153C, YLR291C, YLR312W-
A, YLR344W, YLR372W, YLR432W, YML009C, YML022W, YML086C, YML110C, YMR121C, YMR158W, YMR297W, YMR307W, YNCC0014W, YNCD0005C, YNCD0023C, YNCE0021C, YNCF0005C, YNCG0012W, YNCG0025C, YNCP0017C, YNCQ0024C, YNL010W, YNL071W, YNL085W, YNL130C, YNL185C, YNL209W, YNL247W, YNL302C, YNL306W, YOL064C, YOL096C, YOL097C, YOL139C, YOR046C, YOR074C, YOR133W, YOR168W, YOR187W, YOR198C, YOR224C, YOR230W, YOR312C, YOR323C, YPL028W, YPL131W, YPL237W, YPL240C, YPL252C, YPL273W, YPR033C, YPR110C, YPR187W GO:0009058

mitochondrial transport 0.003286465029409207
YAL005C, YGR082W, YHR005C-
A, YJL104W, YKL084W, YLR008C, YLR259C, YMR203W, YNL070W, YNL121C, YNL131W, YNR017W, YOR020C, YOR045W, YPL063W, YPL240C GO:0006839

mitochondrial transmembrane transport 0.0033280963909424984
YBR085W, YGR082W, YJL104W, YKL084W, YLR008C, YLR259C, YMR203W, YNL070W, YNL121C, YNL131W, YNR017W, YOR020C, YOR045W, YPL063W GO:1990542

organic acid metabolic process 0.0035433865022546454
YBR121C, YBR177C, YBR221C, YBR263W, YCL018W, YCR034W, YCR053W, YDL015C, YDR023W, YDR127W, YDR321W, YER090W, YER178W, YFL018C, YFL022C, YGL105W, YGL148W, YGL245W, YGR094W, YGR124W, YHR183W, YIL078W, YIL094C, YJL052W, YJL088W, YJR139C, YKL192C, YLL018C, YLR153C, YLR372W, YML004C, YNL010W, YNL071W, YNL247W, YOL064C, YOL097C, YOR168W, YOR317W, YOR323C, YPL028W, YPL273W, YPR004C, YPR033C
GO:0006082

organic substance biosynthetic process 0.0037648447381382457

YAL003W, YAL005C, YAL023C, YAL035W, YBL032W, YBL092W, YBR029C, YBR048W, YBR088C, YBR118W, YBR121C, YBR143C, YBR154C, YBR177C, YBR205W, YBR221C, YBR263W, YCL018W, YCL047C, YCL050C, YCR003W, YCR034W, YCR053W, YDL015C, YDL084W, YDL126C, YDL136W, YDL185W, YDR023W, YDR123C, YDR127W, YDR174W, YDR194C, YDR302W, YDR385W, YDR429C, YDR450W, YEL002C, YEL050C, YER003C, YER007C-A, YER074W, YER090W, YER177W, YER178W, YFL018C, YFL022C, YGL031C, YGL068W, YGL105W, YGL148W, YGL245W, YGR094W, YGR124W, YGR157W, YGR180C, YGR285C, YHL003C, YHL034C, YHR064C, YHR123W, YHR147C, YIL078W, YIL094C, YJL034W, YJL052W, YJL088W, YJL134W, YJL167W, YJR101W, YJR104C, YJR105W, YJR139C, YKL006W, YKL035W, YKL056C, YKL058W, YKL067W, YKL109W, YKL117W, YKL127W, YKL138C, YKL192C, YKR057W, YKR059W, YLL018C, YLR069C, YLR150W, YLR153C, YLR291C, YLR312W-A, YLR344W, YLR372W, YLR432W, YML009C, YML022W, YML086C, YML110C, YMR121C, YMR158W, YMR297W, YMR307W, YNCC0014W, YNCD0005C, YNCD0023C, YNCE0021C, YNCF0005C, YNCG0012W, YNCG0025C, YNCP0017C, YNCQ0024C, YNL010W, YNL071W, YNL085W, YNL130C, YNL185C, YNL209W, YNL247W, YNL302C, YNL306W, YOL064C, YOL096C, YOL097C, YOL139C, YOR046C, YOR074C, YOR133W, YOR168W, YOR187W, YOR198C, YOR224C, YOR230W, YOR312C, YOR323C, YPL028W, YPL131W, YPL237W, YPL240C, YPL252C, YPL273W, YPR033C, YPR110C, YPR187W GO:1901576

intracellular protein transport 0.012311050224532372

YAL005C, YAL007C, YAL023C, YBR164C, YDL100C, YDL126C, YDL192W, YEL027W, YER048C, YGL206C, YGR082W, YGR167W, YHR005C-A, YHR193C, YJL034W, YJL104W, YJR074W, YKL084W, YKR001C, YLL024C, YLR008C, YLR259C, YLR262C, YLR301W, YML001W, YML012W, YMR203W, YNL055C, YNL070W, YNL121C, YNL131W, YNR017W, YOL129W, YOR007C, YOR020C, YOR045W, YOR164C, YOR230W, YPL037C, YPL063W, YPL218W, YPL240C GO:0006886

organonitrogen compound metabolic process 0.012703119392504354

YAL003W, YAL005C, YAL023C, YAL035W, YBL032W, YBL092W, YBR029C, YBR048W, YBR092C, YBR118W, YBR121C, YBR127C, YBR143C, YBR205W, YBR221C, YBR263W, YBR286W, YCL018W, YCL043C, YCL047C, YCL050C, YCR003W, YCR034W, YCR053W, YDL126C, YDL136W, YDL185W, YDL227C, YDR023W, YDR127W, YDR194C, YDR302W, YDR321W, YDR385W, YDR429C, YDR435C, YDR450W, YEL002C, YEL050C, YER003C, YER007C-A, YER074W, YER090W, YER177W, YER178W, YFL018C, YFL022C, YFR044C, YGL031C, YGL068W, YGL105W, YGL148W, YGL245W, YGR094W, YGR124W, YGR132C, YGR157W, YGR231C, YGR285C, YHL003C, YHL034C, YHR064C, YHR123W, YHR147C, YHR183W, YIL078W, YIL094C, YJL034W, YJL052W, YJL088W, YJL134W, YJL172W, YJR070C, YJR101W, YJR105W, YJR139C, YKL006W, YKL056C, YKL067W, YKL138C, YKR057W, YKR059W, YLL018C, YLL024C, YLR069C, YLR150W, YLR153C, YLR216C, YLR259C, YLR285W, YLR291C, YLR312W-A, YLR344W, YLR354C, YLR372W, YLR432W, YML009C, YML022W, YMR121C, YMR158W, YMR199W, YMR297W, YNCC0014W, YNCD0005C, YNCD0023C, YNCE0021C, YNCF0005C, YNCG0012W, YNCG0025C, YNCP0017C, YNCQ0024C, YNL010W, YNL055C, YNL071W, YNL085W, YNL130C, YNL185C, YNL209W, YNL247W, YNL302C, YNL306W, YOL064C, YOL097C, YOL139C, YOR046C, YOR074C, YOR133W, YOR168W, YOR187W, YOR198C, YOR230W, YOR253W, YOR312C, YOR317W, YOR323C, YPL091W, YPL106C, YPL131W, YPL237W, YPL240C, YPL252C, YPL273W, YPR033C GO:1901564

'de novo' protein folding 0.021293802648641405
YAL005C, YGR285C, YHR064C, YJL034W, YLL024C, YLR259C, YNL209W, YOR020C, Y
PL240C GO:0006458

amino acid metabolic process 0.02760071768104744
YBR121C, YBR263W, YCL018W, YCR053W, YDR023W, YDR127W, YDR321W, YER090W, Y
FL018C, YFL022C, YGL105W, YGL148W, YGL245W, YGR094W, YGR124W, YIL078W, YIL094C
, YJL088W, YJR139C, YLL018C, YNL010W, YNL247W, YOL064C, YOL097C, YOR168W, YOR32
3C, YPL273W, YPR033C GO:0006520

GO Enrichment Biological process *rrp40-W195R* alone DOWN targets

cytoplasmic translation 3.6690933386791033e-8 YBR084C-
A, YDR500C, YGL076C, YGL135W, YGR027C, YGR214W, YHL001W, YHR021C, YHR203C, YJL1
36C, YJL177W, YLR048W, YLR340W, YML026C, YML091C, YMR142C, YMR194W, YNL067W, YO
L040C, YOL121C, YOL127W, YOR369C GO:0002181

ribosomal small subunit assembly 0.04802281732085983
YGR214W, YHR021C, YLR048W, YOL040C, YOL121C, YOR369C GO:0000028

GO Enrichment Biological process *rrp46-L191H* alone DOWN targets

translational elongation 0.000056978333684954345 Q0140, YDL133C-
A, YER131W, YKL170W, YLR388W, YNCA0002W, YNCA0006C, YNCB0012W, YNCB0013W, YNCC
0013W, YNCD0017W, YNCD0031C, YNCE0026W, YNCF0007W, YNCG0008W, YNCH0013C, YNCI
0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL0037W, YNCL0045W, YNCM
0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO0031W, YNCP0009W, YNCP
0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ
0023W GO:0006414

translation 0.0012744880868182824 Q0140, YBR120C, YDL133C-
A, YER131W, YKL170W, YLR388W, YML129C, YNCA0002W, YNCA0006C, YNCB0012W, YNCB00
13W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG00
08W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL00
37W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO00
31W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ00
15W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W GO:0006412

peptide biosynthetic process 0.0014672635489997186
Q0140, YBR120C, YDL133C-
A, YER131W, YKL170W, YLR388W, YML129C, YNCA0002W, YNCA0006C, YNCB0012W, YNCB00
13W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG00
08W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL00
37W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO00

31W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W GO:0043043

peptide metabolic process 0.0026380191463989773

Q0140, YBR120C, YDL133C-

A, YER131W, YKL170W, YLR388W, YML129C, YNCA0002W, YNCA0006C, YNCB0012W, YNCB0013W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG0008W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL0037W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO0031W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W GO:0006518

amide metabolic process 0.0029610860679342007

Q0140, YBR120C, YBR208C, YDL133C-

A, YER131W, YKL170W, YLR350W, YLR388W, YML129C, YMR246W, YNCA0002W, YNCA0006C, YNCB0012W, YNCB0013W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG0008W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL0037W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO0031W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W GO:0043603

cellular macromolecule biosynthetic process 0.006971981984920854

Q0140, YBR120C, YDL133C-

A, YER054C, YER131W, YKL170W, YLR388W, YML129C, YMR105C, YMR311C, YNCA0002W, YNCA0006C, YNCB0012W, YNCB0013W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG0008W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL0037W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO0031W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W, YPR184W GO:0034645

amide biosynthetic process 0.009119088214840478

Q0140, YBR120C, YDL133C-

A, YER131W, YKL170W, YLR388W, YML129C, YNCA0002W, YNCA0006C, YNCB0012W, YNCB0013W, YNCC0013W, YNCD0017W, YNCD0031C, YNCE0006W, YNCE0026W, YNCF0007W, YNCG0008W, YNCH0013C, YNCI0006W, YNCK0006C, YNCK0016W, YNCL0034W, YNCL0035C, YNCL0037W, YNCL0045W, YNCM0025C, YNCN0020C, YNCO0008W, YNCO0012C, YNCO0020C, YNCO0031W, YNCP0009W, YNCP0022W, YNCQ0002W, YNCQ0009W, YNCQ0011W, YNCQ0012W, YNCQ0015W, YNCQ0021W, YNCQ0023W, YOR173W, YPL203W GO:0043604

GO Enrichment Biological process *rrp40-W195R* alone UP targets

gene expression 4.411272689542685e-12

YBR141C, YBR152W, YCR047C, YDR163W, YDR397C, YGL222C, YGR129W, YGR169C-A, YGR251W, YGR280C, YIL019W, YIL150C, YIR015W, YJL098W, YKL082C, YKR060W, YML1

13W, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0015C, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCH0012W, YNCI0008C, YNCJ0003C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0006W, YNCL0007W, YNCL0033C, YNCL0041C, YNCL0042C, YNCL0045W, YNCM0009C, YNCM0012W, YNCM0021C, YNCM0037W, YNCN0001W, YNCO0020C, YNCO0024C, YNCO0026W, YNCO0031W, YNCP0003W, YNCP0013C, YNCP0014C, YNCP0022W, YOR006C, YOR047C, YOR078W, YOR308C, YPL240C, YPL249C-A, YPR152C GO:0010467

rRNA modification 1.5810317423793532e-8
YBR141C, YCR047C, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C GO:0000154

translational elongation 3.9636450872967826e-7 YGR169C-A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022W, YPL249C-A GO:0006414

rRNA methylation 0.0000010379300960873824
YBR141C, YCR047C, YNCG0015C, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCP0003W, YNCP0013C, YNCP0014C GO:0031167

cellular nitrogen compound metabolic process 0.0000021207707688107566
YBR141C, YBR152W, YCR047C, YDR163W, YDR397C, YGL222C, YGR129W, YGR169C-A, YGR251W, YGR280C, YIL019W, YIL150C, YIR015W, YJL098W, YKL082C, YKR060W, YML113W, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0015C, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCH0012W, YNCI0008C, YNCJ0003C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0006W, YNCL0007W, YNCL0033C, YNCL0041C, YNCL0042C, YNCL0045W, YNCM0009C, YNCM0012W, YNCM0021C, YNCM0037W, YNCN0001W, YNCO0020C, YNCO0024C, YNCO0026W, YNCO0031W, YNCP0003W, YNCP0013C, YNCP0014C, YNCP0022W, YOR006C, YOR047C, YOR078W, YOR308C, YPL240C, YPL249C-A, YPR152C GO:0034641

RNA processing 0.00001088189513044786
YBR141C, YBR152W, YCR047C, YDR163W, YGL222C, YGR129W, YGR251W, YGR280C, YIL019W, YIR015W, YJL098W, YKL082C, YKR060W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0006W, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR308C, YPR152C GO:0006396

rRNA processing 0.000012436378032023
YBR141C, YCR047C, YGR251W, YGR280C, YIL019W, YKL082C, YKR060W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR308C GO:0006364

RNA modification 0.000015568667885016763

YBR141C, YCR047C, YJL098W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C GO:0009451

ribonucleoprotein complex biogenesis 0.00009382718604798154

YBR141C, YBR152W, YCR047C, YGR251W, YGR280C, YIL019W, YKL082C, YKR060W, YMR269W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0006W, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR308C, YPL240C GO:0022613

ncRNA processing 0.00011118634493951462

YBR141C, YCR047C, YGR251W, YGR280C, YIL019W, YIR015W, YJL098W, YKL082C, YKR060W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR308C GO:0034470

RNA methylation 0.00011256077965893588

YBR141C, YCR047C, YNCG0015C, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCP0003W, YNCP0013C, YNCP0014C GO:0001510

rRNA metabolic process 0.0001454623838078644

YBR141C, YCR047C, YGR251W, YGR280C, YIL019W, YKL082C, YKR060W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR308C GO:0016072

translation 0.00017963278942383952 YGL222C, YGR169C-

A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022W, YPL249C-A GO:0006412

peptide biosynthetic process 0.00020198301085736655

YGL222C, YGR169C-A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022W, YPL249C-A GO:0043043

peptide metabolic process 0.00032975813314502163 YGL222C, YGR169C-

A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022W, YPL249C-A GO:0006518

ribosome biogenesis 0.0003332034345959769
YBR141C, YCR047C, YGR251W, YGR280C, YIL019W, YKL082C, YKR060W, YMR269W, Y
NCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNCN0001W, Y
NCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR078W, YOR30
8C GO:0042254

amide biosynthetic process 0.0009386737784076849 YGL222C, YGR169C-
A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027
C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046
W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033
C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022
W, YPL249C-A GO:0043604

macromolecule metabolic process 0.002243850712947081
YBR141C, YBR152W, YCR047C, YDR163W, YDR397C, YGL222C, YGR129W, YGR169C-
A, YGR188C, YGR251W, YGR280C, YIL019W, YIL150C, YIR015W, YJL098W, YKL082C, YKR0
60W, YML113W, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017
W, YNCD0027C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0015C, YNCG0021
W, YNCG0040C, YNCG0046W, YNCH0006C, YNCH0012W, YNCI0008C, YNCJ0003C, YNCJ0008
W, YNCK0006C, YNCK0014W, YNCL0006W, YNCL0007W, YNCL0033C, YNCL0041C, YNCL0042
C, YNCL0045W, YNCM0009C, YNCM0012W, YNCM0021C, YNCM0037W, YNCN0001W, YNCO0020
C, YNCO0024C, YNCO0026W, YNCO0031W, YNCP0003W, YNCP0013C, YNCP0014C, YNCP0022
W, YOR006C, YOR047C, YOR078W, YOR308C, YPL240C, YPL249C-A, YPR152C GO:0043170

amide metabolic process 0.004006389199241653 YGL222C, YGR169C-
A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027
C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046
W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033
C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022
W, YPL249C-A GO:0043603

macromolecule methylation 0.005269117671700851
YBR141C, YCR047C, YNCG0015C, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W
, YNCN0001W, YNCP0003W, YNCP0013C, YNCP0014C GO:0043414

ncRNA metabolic process 0.006571076635598721
YBR141C, YCR047C, YGR251W, YGR280C, YIL019W, YIR015W, YJL098W, YKL082C, Y
KR060W, YNCG0015C, YNCH0012W, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W, YNC
N0001W, YNCO0024C, YNCO0026W, YNCP0003W, YNCP0013C, YNCP0014C, YOR006C, YOR07
8W, YOR308C GO:0034660

cellular macromolecule biosynthetic process 0.015737008635486785
YGL222C, YGR169C-
A, YNCA0005W, YNCB0001W, YNCC0010C, YNCD0007C, YNCD0013C, YNCD0017W, YNCD0027
C, YNCE0016C, YNCE0018W, YNCF0009C, YNCG0004W, YNCG0021W, YNCG0040C, YNCG0046
W, YNCH0006C, YNCI0008C, YNCJ0008W, YNCK0006C, YNCK0014W, YNCL0007W, YNCL0033
C, YNCL0045W, YNCM0009C, YNCM0021C, YNCM0037W, YNCO0020C, YNCO0031W, YNCP0022
W, YPL249C-A GO:0034645

methylation 0.02266330888556486
YBR141C, YCR047C, YNCG0015C, YNCJ0003C, YNCL0041C, YNCL0042C, YNCM0012W
, YNCN0001W, YNCP0003W, YNCP0013C, YNCP0014C GO:0032259

GO Enrichment Biological process *rrp46-L191H* alone UP targets

ncRNA processing 3.4156696856523035e-10
YBR247C, YCL054W, YDR021W, YDR299W, YDR312W, YDR324C, YDR398W, YDR478W, Y
GL111W, YGL169W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W, YJL010C, YJL035C
, YJL069C, YJL191W, YKL110C, YKL208W, YKR024C, YLL035W, YLR196W, YML080W, YMR04
9C, YMR093W, YMR259C, YNCE0002W, YNCG0014C, YNCM0013W, YNL112W, YNL182C, YNR01
5W, YOL080C, YOR294W, YPL012W, YPL030W, YPL183C, YPR137W, YPR144C GO:0034470

ribosome biogenesis 3.847467859594019e-9
YBR247C, YCL054W, YCR072C, YDL063C, YDR021W, YDR101C, YDR299W, YDR312W, Y
DR324C, YDR398W, YDR478W, YGL111W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W
, YJL010C, YJL069C, YJL122W, YJL191W, YJR063W, YKL208W, YKR024C, YLL035W, YLR19
6W, YLR336C, YMR049C, YMR093W, YNCE0002W, YNCG0014C, YNCM0013W, YNL112W, YNL18
2C, YOL080C, YOR294W, YPL012W, YPR137W, YPR144C GO:0042254

ncRNA metabolic process 5.643236704264575e-9
YBR247C, YCL054W, YDR021W, YDR299W, YDR312W, YDR324C, YDR398W, YDR478W, Y
GL111W, YGL169W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W, YJL010C, YJL035C
, YJL069C, YJL191W, YJR063W, YKL110C, YKL208W, YKR024C, YLL035W, YLR196W, YML04
3C, YML080W, YMR049C, YMR093W, YMR259C, YNCE0002W, YNCG0014C, YNCM0013W, YNL11
2W, YNL182C, YNR003C, YNR015W, YOL080C, YOR294W, YPL012W, YPL030W, YPL183C, YPR
137W, YPR144C GO:0034660

rRNA processing 2.469310981076525e-8
YBR247C, YCL054W, YDR021W, YDR299W, YDR312W, YDR324C, YDR398W, YDR478W, Y
GL111W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W, YJL010C, YJL069C, YJL191W
, YKL208W, YKR024C, YLL035W, YLR196W, YMR049C, YMR093W, YNCE0002W, YNCG0014C, Y
NCM0013W, YNL112W, YNL182C, YOL080C, YOR294W, YPL012W, YPR137W, YPR144C
GO:0006364

rRNA metabolic process 4.085552645506523e-8
YBR247C, YCL054W, YDR021W, YDR299W, YDR312W, YDR324C, YDR398W, YDR478W, Y
GL111W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W, YJL010C, YJL069C, YJL191W
, YJR063W, YKL208W, YKR024C, YLL035W, YLR196W, YML043C, YMR049C, YMR093W, YNCE0
002W, YNCG0014C, YNCM0013W, YNL112W, YNL182C, YOL080C, YOR294W, YPL012W, YPR13
7W, YPR144C GO:0016072

ribonucleoprotein complex biogenesis 1.9611651522586821e-7
YBR247C, YCL054W, YCR072C, YDL063C, YDR021W, YDR101C, YDR299W, YDR312W, Y
DR324C, YDR398W, YDR478W, YGL111W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W
, YIL104C, YJL010C, YJL069C, YJL122W, YJL191W, YJR063W, YKL208W, YKR024C, YLL03

5W, YLR196W, YLR336C, YMR049C, YMR093W, YNCE0002W, YNCG0014C, YNCM0013W, YNL112W, YNL182C, YOL080C, YOR294W, YPL012W, YPR137W, YPR144C GO:0022613

RNA processing 0.00000353891450876183

YBR247C, YCL054W, YCR063W, YDR021W, YDR299W, YDR312W, YDR324C, YDR398W, YDR478W, YGL111W, YGL169W, YGR095C, YGR159C, YHR066W, YHR085W, YHR196W, YJL010C, YJL035C, YJL069C, YJL191W, YKL110C, YKL208W, YKR024C, YLL035W, YLR196W, YML080W, YMR049C, YMR093W, YMR259C, YNCE0002W, YNCG0014C, YNCM0013W, YNL112W, YNL182C, YNR015W, YOL080C, YOR294W, YPL012W, YPL030W, YPL183C, YPR137W, YPR144C
GO:0006396

maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) 0.0007016505141477728

YBR247C, YDR021W, YDR299W, YDR324C, YDR398W, YHR196W, YJL010C, YJL069C, YJL191W, YMR093W, YOR294W, YPL012W, YPR137W, YPR144C GO:0000462

ribosomal small subunit biogenesis 0.000934301433980673

YBR247C, YDR021W, YDR299W, YDR324C, YDR398W, YGR159C, YHR196W, YJL010C, YJL069C, YJL191W, YLR336C, YMR093W, YOR294W, YPL012W, YPR137W, YPR144C
GO:0042274

maturation of SSU-rRNA 0.0032566333582539535

YBR247C, YDR021W, YDR299W, YDR324C, YDR398W, YHR196W, YJL010C, YJL069C, YJL191W, YMR093W, YOR294W, YPL012W, YPR137W, YPR144C GO:0030490

ribosomal large subunit biogenesis 0.004396160899671331

YCL054W, YCR072C, YDL063C, YDR312W, YGL111W, YHR066W, YHR085W, YJL122W, YKR024C, YLL035W, YMR049C, YNL182C, YOL080C, YOR294W GO:0042273

GO Enrichment Biological process *rrp40-W195R* intersection with *rrp4-G226D* DOWN targets

cytoplasmic translation 3.7612510240247765e-29

YBL027W, YBR191W, YDL082W, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YIL018W, YIL133C, YJL190C, YJR094W-A, YJR145C, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL220W GO:0002181

organonitrogen compound biosynthetic process 4.098191531886011e-17

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR226W, YDR354W, YDR418W, YDR447C, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL125W, YGL147C, YGR034W, YGR085C, YGR149W, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR073C, YJR094W-A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR0

75W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:1901566

amide biosynthetic process 1.3353787705481637e-14

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR094W-

A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W

GO:0043604

amide metabolic process 3.279135436374408e-14

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR094W-

A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W

GO:0043603

peptide metabolic process 8.974366878301551e-14

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR094W-

A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0006518

translation 2.0552730111936557e-13

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR094W-

A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0006412

peptide biosynthetic process 2.514513723664123e-13

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR094W-

A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR4

41C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0043043

translational elongation 6.314549568398182e-12

YBL027W, YBR191W, YDL082W, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YIL018W, YIL133C, YJL190C, YJR094W-A, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL220W GO:0006414

cellular macromolecule biosynthetic process 9.639291628850999e-11

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR094W-A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0034645

organonitrogen compound metabolic process 1.2185721637959406e-10

YAL038W, YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YDR502C, YEL040W, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL125W, YGL147C, YGR034W, YGR052W, YGR085C, YGR149W, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR073C, YJR094W-A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL015W, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL004C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:1901564

organic substance biosynthetic process 2.8777630754599182e-8

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YDR502C, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL125W, YGL147C, YGR034W, YGR085C, YGR149W, YHR010W, YHR019C, YIL018W, YIL124W, YIL133C, YJL153C, YJL190C, YJR073C, YJR094W-A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:1901576

cellular biosynthetic process 3.232474027036089e-8

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YDR502C, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL125W, YGL147C, YGR034W, YGR085C, YGR149W, YHR010W, YHR019C, YIL018W, YIL1

24W, YIL133C, YJL153C, YJL190C, YJR073C, YJR094W-
A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR0
43C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YM
R242C, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W,
YPL220W, YPR069C, YPR132W, YPR145W GO:0044249

biosynthetic process 4.2241456923363024e-8

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, Y
DR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YDR502C, YEL054C, YER023W
, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-
A, YGL125W, YGL147C, YGR034W, YGR085C, YGR149W, YHR010W, YHR019C, YIL018W, YIL1
24W, YIL133C, YJL153C, YJL190C, YJR073C, YJR094W-
A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR0
43C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YM
L073C, YMR242C, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C,
YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:0009058

ribosome biogenesis 1.3110394419135715e-7

YBL072C, YBR189W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, Y
FR031C-
A, YGR034W, YGR085C, YIL133C, YJR123W, YKL180W, YKR094C, YLL045C, YLR075W, YLR4
41C, YLR448W, YML024W, YML063W, YML073C, YNCM0034C, YNCN0019C, YNL096C, YOR096
W, YOR234C, YPL090C, YPL220W, YPR132W GO:0042254

ribonucleoprotein complex biogenesis 0.000009060391633841293

YBL072C, YBR189W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, Y
FR031C-
A, YGR034W, YGR085C, YIL133C, YJR123W, YKL180W, YKR094C, YLL045C, YLR075W, YLR4
41C, YLR448W, YML024W, YML063W, YML073C, YNCM0034C, YNCN0019C, YNL096C, YOR096
W, YOR234C, YPL090C, YPL220W, YPR132W GO:0022613

cellular nitrogen compound biosynthetic process

0.000012259522080692909

YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL066W, YDL082W, YDL083C, Y
DR025W, YDR226W, YDR354W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C
-A, YFL045C, YFR031C-
A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR0
94W-
A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR0
43C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YM
R242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W,
YPR069C, YPR132W, YPR145W GO:0044271

ribosome assembly 0.000261332470694884

YDR025W, YDR418W, YDR447C, YEL054C, YGR085C, YJR123W, YKR094C, YLR075W, Y
LR448W, YML024W, YML073C GO:0042255

ribosomal small subunit biogenesis 0.00033166321849001856
YBL072C, YBR189W, YDL083C, YDR025W, YDR447C, YER102W, YJR123W, YLR441C, YML024W, YML063W, YNL096C, YOR096W, YPL090C, YPR132W GO:0042274

ribosomal large subunit biogenesis 0.00038810259077208233
YDR418W, YEL054C, YFR031C-
A, YGR034W, YGR085C, YIL133C, YKL180W, YKR094C, YLL045C, YLR075W, YLR448W, YML073C, YOR234C GO:0042273

protein metabolic process 0.00047195292625167864
YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YDL082W, YDL083C, YDR025W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-
A, YGL147C, YGR034W, YGR052W, YGR085C, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR094W-
A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL015W, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL004C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0019538

cellular macromolecule metabolic process 0.009209834745154058
Q0160, YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YCR028C-
A, YDL082W, YDL083C, YDR025W, YDR226W, YDR418W, YDR447C, YEL054C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-
A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR045C, YJR094W-
A, YJR123W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR075W, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W GO:0044260

nitrogen compound metabolic process 0.01448490168901606
Q0160, YAL038W, YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YCR028C-
A, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YDR502C, YEL040W, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-
A, YGL125W, YGL147C, YGR034W, YGR052W, YGR085C, YGR149W, YHR010W, YHR019C, YHR047C, YIL018W, YIL133C, YJL190C, YJR045C, YJR073C, YJR094W-
A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNCM0034C, YNCN0019C, YNL015W, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL004C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:0006807

ribosomal large subunit assembly 0.016504696397229995
YDR418W, YEL054C, YGR085C, YKR094C, YLR075W, YLR448W, YML073C
GO:0000027

cellular nitrogen compound metabolic process 0.017196664727524826
Q0160, YAL038W, YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YCR028C-

A, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR418W, YDR447C, YEL054C, YER043C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL147C, YGR034W, YGR085C, YHR010W, YHR019C, YIL018W, YIL133C, YJL190C, YJR045C, YJR094W-
A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNCM0034C, YNCN0019C, YNL096C, YOL120C, YOR096W, YOR234C, YPL079W, YPL090C, YPL160W, YPL220W, YPR069C, YPR132W, YPR145W GO:0034641

primary metabolic process 0.03530585522621974

Q0160, YAL038W, YBL027W, YBL072C, YBL076C, YBR189W, YBR191W, YCR028C-A, YDL066W, YDL082W, YDL083C, YDR025W, YDR050C, YDR226W, YDR354W, YDR368W, YDR418W, YDR447C, YDR502C, YEL040W, YEL054C, YER023W, YER026C, YER043C, YER052C, YER102W, YER117W, YFL034C-A, YFL045C, YFR031C-A, YGL125W, YGL147C, YGR034W, YGR052W, YGR085C, YGR149W, YHR010W, YHR019C, YHR047C, YIL018W, YIL124W, YIL133C, YJL153C, YJL190C, YJR045C, YJR073C, YJR094W-A, YJR109C, YJR123W, YJR133W, YJR145C, YKL081W, YKL180W, YKR094C, YLL045C, YLR043C, YLR044C, YLR075W, YLR231C, YLR333C, YLR441C, YLR448W, YML024W, YML063W, YML073C, YMR242C, YNCM0034C, YNCN0019C, YNL015W, YNL096C, YNL169C, YOL120C, YOR096W, YOR234C, YPL004C, YPL079W, YPL090C, YPL160W, YPL220W, YPR132W, YPR145W
GO:0044238

non-membrane-bounded organelle assembly 0.0397240666921655

YDR025W, YDR418W, YDR447C, YEL054C, YGR085C, YJR123W, YKR094C, YLR075W, YLR448W, YML024W, YML073C, YNL169C GO:0140694

GO Enrichment Biological process *rrp46-L191H* intersection with *rrp4-G226D* DOWN targets

translational elongation 0.0000066136211740309875

YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0006414

cellular macromolecule biosynthetic process 0.0009347304769258927

YER001W, YIL051C, YLR406C, YMR215W, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C
GO:0034645

translation 0.0011926707779502603

YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0006412

peptide biosynthetic process 0.0012910478172846421
YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W
, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W
, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0043043

peptide metabolic process 0.0017997833532501552
YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W
, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W
, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0006518

amide biosynthetic process 0.0036726033978837624
YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W
, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W
, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0043604

mitochondrial translational elongation 0.003971961509304703
YNCQ0004W, YNCQ0005W, YNCQ0020W, YNCQ0025W, YNCQ0026W GO:0070125

organonitrogen compound biosynthetic process 0.0089037576929183
YER001W, YGL009C, YHR137W, YIL051C, YJR025C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:1901566

amide metabolic process 0.009973896074848837
YIL051C, YLR406C, YNCG0011W, YNCG0020C, YNCK0002C, YNCK0003W, YNCL0047W
, YNCN0004W, YNCO0013C, YNCO0021C, YNCP0008C, YNCQ0004W, YNCQ0005W, YNCQ0020W
, YNCQ0025W, YNCQ0026W, YNL252C, YPR100W, YPR102C GO:0043603

GO Enrichment Biological process *rrp40-W195R* intersection with *rrp46-L191H* UP targets

rRNA metabolic process 0.0015498741592715496
YER127W, YFL023W, YHR040W, YHR169W, YKL078W, YLR063W, YNCK0009W, YNCL0005C, YNCM0017W, YNL023C, YNL124W, YNL299W, YOR287C GO:0016072

ncRNA processing 0.009411132247541193
YER127W, YFL023W, YHR040W, YHR169W, YKL078W, YLR063W, YNCK0009W, YNCL0005C, YNCM0017W, YNL124W, YNL299W, YOL141W, YOR287C GO:0034470

ncRNA metabolic process 0.017725959980186123
YER127W, YFL023W, YHR040W, YHR169W, YKL078W, YLR063W, YNCK0009W, YNCL0005C, YNCM0017W, YNL023C, YNL124W, YNL299W, YOL141W, YOR287C GO:0034660

rRNA processing 0.02338880898778791
YER127W, YFL023W, YHR040W, YHR169W, YKL078W, YLR063W, YNCK0009W, YNCL0005C, YNCM0017W, YNL124W, YOR287C GO:0006364

ribosome biogenesis 0.04802869746543283
YBR267W, YER127W, YFL023W, YHR040W, YHR169W, YKL078W, YLR063W, YNCK0009W, YNCL0005C, YNCM0017W, YNL124W, YOR287C GO:0042254

GO Enrichment Biological process *rrp46-L191H* intersection with *rrp4-G226D* UP targets

telomere maintenance via recombination 0.002652724434062646
YDR545W, YER190W, YGR296W, YLR467W, YML133C, YOR396W, YPL283C
GO:0000722

mitotic recombination 0.007671662555864012
YDR545W, YER190W, YGR296W, YLR467W, YML133C, YOR396W, YPL017C, YPL283C
GO:0006312

Supplemental Documentation S4. Full Gene Ontology (GO) Analyses on Human homologs

Gene list: Identified Human homologs of *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* DOWN targets

Gene > Symbol	Gene > Name	Gene > Primary Identifier
AARSD1	alanyl-tRNA synthetase domain containing 1	80755
ADSL	adenylosuccinate lyase	158
ALDH1A1	aldehyde dehydrogenase 1 family member A1	216
ALDH1A2	aldehyde dehydrogenase 1 family member A2	8854
ALDH1A3	aldehyde dehydrogenase 1 family member A3	220
ALDH1B1	aldehyde dehydrogenase 1 family member B1	219
ALDH1L1	aldehyde dehydrogenase 1 family member L1	10840
ALDH1L2	aldehyde dehydrogenase 1 family member L2	160428
ALDH2	aldehyde dehydrogenase 2 family member	217
ASL	argininosuccinate lyase	435

ATIC 5-aminoimidazole-4-carboxamide ribonucleotide
 formyltransferase/IMP cyclohydrolase 471

BCAT1 branched chain amino acid transaminase 1 586

BCAT2 branched chain amino acid transaminase 2 587

CTTN cortactin 2017

DBN1 drebrin 1 1627

DBNL drebrin like 28988

EIF5A eukaryotic translation initiation factor 5A 1984

EIF5A2 eukaryotic translation initiation factor 5A2 56648

EIF5AL1 eukaryotic translation initiation factor 5A like 1
 143244

ENO1 enolase 1 2023

ENO2 enolase 2 2026

ENO3 enolase 3 2027

ENO4 enolase 4 387712

GAPDH glyceraldehyde-3-phosphate dehydrogenase 2597

GAPDHS glyceraldehyde-3-phosphate dehydrogenase, spermatogenic
 26330

GCK glucokinase 2645

GMPPB GDP-mannose pyrophosphorylase B 29925

GPI glucose-6-phosphate isomerase 2821

H2AB1 H2A.B variant histone 1 474382

H2AB3 H2A.B variant histone 3 83740

H2AC1 H2A clustered histone 1 221613

H2AC12 H2A clustered histone 12 85235

H2AC13 H2A clustered histone 13 8329

H2AC14 H2A clustered histone 14 8331

H2AC19 H2A clustered histone 19 723790

H2AC20 H2A clustered histone 20 8338

H2AC21 H2A clustered histone 21 317772

H2AC6 H2A clustered histone 6 8334

H2AC7	H2A clustered histone	7	3013
H2AC8	H2A clustered histone	8	3012
H2AJ	H2A.J histone		55766
H2AL3	H2A.L variant histone	3	115482686
H2AP	H2A.P histone		25763
H2AW	H2A.W histone		92815
H2AX	H2A.X variant histone		3014
HCLS1	hematopoietic cell-specific Lyn substrate	1	3059
HK1	hexokinase	1	3098
HK2	hexokinase	2	3099
HK3	hexokinase	3	3101
HKDC1	hexokinase domain containing	1	80201
MACROH2A1	macroH2A.1 histone		9555
MACROH2A2	macroH2A.2 histone		55506
MDH2	malate dehydrogenase	2	4191
MECR	mitochondrial trans-2-enoyl-CoA reductase		51102
MRAS	muscle RAS oncogene homolog		22808
NAGS	N-acetylglutamate synthase		162417
PGK1	phosphoglycerate kinase	1	5230
PGK2	phosphoglycerate kinase	2	5232
PRKAB1	protein kinase AMP-activated non-catalytic subunit beta	1	5564
PRKAB2	protein kinase AMP-activated non-catalytic subunit beta	2	5565
PSAT1	phosphoserine aminotransferase	1	29968
PYGB	glycogen phosphorylase	B	5834
PYGL	glycogen phosphorylase	L	5836
PYGM	glycogen phosphorylase, muscle associated		5837
QPRT	quinolinate phosphoribosyltransferase		23475
RACK1	receptor for activated C kinase	1	10399
RERG	RAS like estrogen regulated growth inhibitor		85004

RIT1	Ras like without CAAX	1	6016
RIT2	Ras like without CAAX	2	6014
RPL15	ribosomal protein L15		6138
RPL24	ribosomal protein L24		6152
RPL3	ribosomal protein L3		6122
RPL30	ribosomal protein L30		6156
RPL3L	ribosomal protein L3 like		6123
RPL4	ribosomal protein L4		6124
RPS13	ribosomal protein S13		6207
RPS2	ribosomal protein S2		6187
RPS20	ribosomal protein S20		6224
RPS27A	ribosomal protein S27a		6233
RPS27AP5	ribosomal protein S27a pseudogene 5		100271374
RPS29	ribosomal protein S29		6235
RPS3	ribosomal protein S3		6188
RRAS	RAS related		6237
RRAS2	RAS related 2		22800
SHMT1	serine hydroxymethyltransferase 1		6470
SHMT2	serine hydroxymethyltransferase 2		6472
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1		7311
UBB	ubiquitin B		7314
UBC	ubiquitin C		7316

Gene list: Identified Human homologs of *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* UP targets

Gene > Symbol	Gene > Name	Gene > Primary Identifier
ANAPC10	anaphase promoting complex subunit 10	10393
CD2BP2	CD2 cytoplasmic tail binding protein	210421

NAA40	N-alpha-acetyltransferase 40, NatD catalytic subunit	79829
RPL18	ribosomal protein L18	6141
SNAP23	synaptosome associated protein 23	8773
SNAP25	synaptosome associated protein 25	6616
SNAP29	synaptosome associated protein 29	9342
SNAP47	synaptosome associated protein 47	116841
TADA2A	transcriptional adaptor 2A	6871
TADA2B	transcriptional adaptor 2B	93624
TBCD	tubulin folding cofactor D	6904
TOLLIP	toll interacting protein	54472
UBE2V1	ubiquitin conjugating enzyme E2 V1	7335
UBE2V2	ubiquitin conjugating enzyme E2 V2	7336

GO Enrichment Biological process for identified Human homologs of *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* DOWN targets

carboxylic acid metabolic process 2.5559192563708388e-17
10840,160428,162417,2023,2026,2027,216,220,23475,2597,26330,2645,
2821,29968,3098,3099,3101,387712,4191,471,51102,5230,5232,5564,5565,58
6,587,6470,6472,80201,80755,8854 GO:0019752

oxoacid metabolic process 5.782212707588229e-17
10840,160428,162417,2023,2026,2027,216,220,23475,2597,26330,2645,
2821,29968,3098,3099,3101,387712,4191,471,51102,5230,5232,5564,5565,58
6,587,6470,6472,80201,80755,8854 GO:0043436

organic acid metabolic process 1.196879611920814e-16
10840,160428,162417,2023,2026,2027,216,220,23475,2597,26330,2645,
2821,29968,3098,3099,3101,387712,4191,471,51102,5230,5232,5564,5565,58
6,587,6470,6472,80201,80755,8854 GO:0006082

glycolytic process 1.0152884711137568e-14
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0006096

ATP generation from ADP 1.2200003586090464e-14
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0006757

ADP metabolic process 4.8777167394558457e-14
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0046031

carbohydrate catabolic process 6.612308507141743e-14
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,5834,5836,80201 GO:0016052

nucleoside diphosphate phosphorylation 1.7122767299801073e-13
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0006165

nucleotide phosphorylation 1.9872601781131742e-13
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0046939

purine nucleoside diphosphate metabolic process 3.5486491755122165e-
13
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0009135

purine ribonucleoside diphosphate metabolic process
3.5486491755122165e-13
2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
32,80201 GO:0009179

pyruvate metabolic process 5.394357871632931e-13
 2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
 32,80201 GO:0006090

ribonucleoside diphosphate metabolic process 1.200672980316271e-12
 2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
 32,80201 GO:0009185

generation of precursor metabolites and energy 1.3085246322426014e-
 12
 10840,158,160428,2023,2026,2027,217,2597,26330,2645,2821,3098,309
 9,3101,387712,4191,5230,5232,5834,5836,6472,80201GO:0006091

small molecule metabolic process 1.6879959393869294e-12
 10840,158,160428,162417,2023,2026,2027,216,217,219,220,23475,2597
 ,26330,2645,2821,29925,29968,3098,3099,3101,387712,4191,471,51102,5230
 ,5232,5564,5565,586,587,6470,6472,80201,80755,8854 GO:0044281

nucleoside diphosphate metabolic process 9.187625672830392e-12
 2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
 32,80201 GO:0009132

organonitrogen compound biosynthetic process 3.0001824927601076e-
 10
 10399,143244,158,162417,1984,2023,216,217,23475,2597,29925,29968,
 471,56648,586,587,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,62
 33,6235,6470,6472,7311,80755 GO:1901566

translation 1.4113741515761663e-9
 10399,143244,1984,2597,56648,6122,6123,6124,6138,6152,6156,6187,6
 188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0006412

monocarboxylic acid metabolic process 1.5080928589743614e-9
 160428,2023,2026,2027,216,220,2597,26330,2645,2821,3098,3099,3101
 ,387712,51102,5230,5232,5564,5565,80201,8854 GO:0032787

purine-containing compound metabolic process 2.245873000988796e-9
 158,2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,471
 ,5230,5232,6470,6472,80201 GO:0072521

peptide biosynthetic process 3.2794537700616576e-9
 10399,143244,1984,2597,56648,6122,6123,6124,6138,6152,6156,6187,6
 188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0043043

nucleobase-containing small molecule metabolic process
 4.109676745684942e-9
 158,2023,2026,2027,23475,2597,26330,2645,2821,29925,3098,3099,310
 1,387712,471,5230,5232,6470,6472,80201GO:0055086

ATP metabolic process 8.0311569447065e-9
 2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,5230,52
 32,6472,80201 GO:0046034

cytoplasmic translation 1.1133082687558695e-8
6122,6124,6138,6152,6156,6187,6188,6207,6224,6233,6235,7311
GO:0002181

ribose phosphate metabolic process 1.1845052412768421e-8
158,2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,471
,5230,5232,5836,80201 GO:0019693

amide biosynthetic process 1.384322424302684e-8
10399,143244,162417,1984,2597,56648,6122,6123,6124,6138,6152,6156
,6187,6188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0043604

nucleotide metabolic process 3.480075038079054e-8
158,2023,2026,2027,23475,2597,26330,2645,2821,3098,3099,3101,3877
12,471,5230,5232,6470,80201 GO:0009117

nucleoside phosphate metabolic process 4.46238633478257e-8
158,2023,2026,2027,23475,2597,26330,2645,2821,3098,3099,3101,3877
12,471,5230,5232,6470,80201 GO:0006753

NADH regeneration 5.151507778499602e-8
2023,2026,2645,3098,3099,3101 GO:0006735

canonical glycolysis 5.151507778499602e-8
2023,2026,2645,3098,3099,3101 GO:0061621

glucose catabolic process to pyruvate 5.151507778499602e-8
2023,2026,2645,3098,3099,3101 GO:0061718

purine ribonucleotide metabolic process 5.420950278762323e-8
158,2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,471
,5230,5232,80201 GO:0009150

carbohydrate metabolic process 1.0031714236545244e-7
2023,2026,2027,217,219,2597,26330,2645,2821,3098,3099,3101,387712
,4191,5230,5232,5834,5836,80201 GO:0005975

ribonucleotide metabolic process 1.1077141667971644e-7
158,2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,471
,5230,5232,80201 GO:0009259

purine nucleotide metabolic process 1.6582970415847214e-7
158,2023,2026,2027,2597,26330,2645,2821,3098,3099,3101,387712,471
,5230,5232,80201 GO:0006163

glycolytic process through glucose-6-phosphate 1.8988745986490162e-7
2023,2026,2645,3098,3099,3101 GO:0061620

peptide metabolic process 2.5580633493831806e-7
10399,143244,1984,2597,56648,6122,6123,6124,6138,6152,6156,6187,6
188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0006518

glycolytic process through fructose-6-phosphate 3.310693777925773e-7
2023,2026,2645,3098,3099,3101 GO:0061615

NAD metabolic process 5.21009130762428e-7
2023,2026,23475,2645,3098,3099,3101 GO:0019674

NADH metabolic process 6.842632984296031e-7
2023,2026,2645,3098,3099,3101,4191 GO:0006734

cellular amide metabolic process 0.0000013278833456433102
10399,143244,160428,162417,1984,2597,56648,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0043603

small molecule biosynthetic process 0.0000028339850389975736
162417,2026,216,220,2645,2821,29968,51102,5230,5232,5564,5565,586,587,6470,6472,8854 GO:0044283

glucose catabolic process 0.0000029370219989830507
2023,2026,2645,3098,3099,3101 GO:0006007

hexose metabolic process 0.000007684419345933139
2023,2026,2597,26330,2645,2821,3098,3099,3101,5230,5232,80201 GO:0019318

glucose 6-phosphate metabolic process 0.000010770204347962276
2645,2821,3098,3099,3101,80201 GO:0051156

small molecule catabolic process 0.000013683894599237153
10840,160428,2023,2026,217,219,23475,2645,3098,3099,3101,587,6470,6472 GO:0044282

glucose metabolic process 0.000014380967280704786
2023,2026,2597,26330,2645,2821,3098,3099,3101,5230,5232 GO:0006006

organic substance catabolic process 0.00001442051365683671
10399,10840,160428,2023,2026,2027,216,217,219,23475,2597,26330,2645,2821,3098,3099,3101,387712,5230,5232,5834,5836,587,6233,6470,6472,7311,7314,7316,80201 GO:1901575

monosaccharide metabolic process 0.000016580226941326802
2023,2026,2597,26330,2645,2821,3098,3099,3101,5230,5232,80201 GO:0005996

organonitrogen compound metabolic process 0.0000328707582360573
10399,10840,143244,158,160428,162417,1984,2023,2026,2027,216,217,23475,2597,26330,2645,2821,29925,29968,3059,3098,3099,3101,387712,471,5230,5232,5564,5565,56648,586,587,6014,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,6233,6235,6470,6472,7311,7314,7316,80201,80755,9555 GO:1901564

carboxylic acid biosynthetic process 0.00010645718804131975
 162417,216,220,29968,51102,5564,5565,586,587,6470,6472,8854
 GO:0046394

hexose catabolic process 0.00011394915009436024
 2023,2026,2645,3098,3099,3101 GO:0019320

organic acid biosynthetic process 0.0001149568279215437
 162417,216,220,29968,51102,5564,5565,586,587,6470,6472,8854
 GO:0016053

carbohydrate derivative metabolic process 0.00014000722481765672
 158,2023,2026,2027,216,2597,26330,2645,2821,29925,3098,3099,3101,
 387712,471,5230,5232,5836,6470,80201 GO:1901135

dicarboxylic acid metabolic process 0.0001594683779688396
 10840,160428,162417,23475,4191,471,6470,6472 GO:0043648

tetrahydrofolate metabolic process 0.00021846297456727155
 10840,160428,471,6470,6472 GO:0046653

monosaccharide catabolic process 0.00023569352833551128
 2023,2026,2645,3098,3099,3101 GO:0046365

organophosphate metabolic process 0.0002364440867358798
 158,2023,2026,2027,23475,2597,26330,2645,2821,3098,3099,3101,3877
 12,471,5230,5232,5836,6470,80201 GO:0019637

catabolic process 0.0003670751680063375
 10399,10840,160428,2017,2023,2026,2027,216,217,219,23475,2597,263
 30,2645,2821,3098,3099,3101,387712,5230,5232,5834,5836,587,6233,6470,6
 472,7311,7314,7316,80201 GO:0009056

cellular nitrogen compound metabolic process
 0.00037010939357173063
 10399,10840,143244,158,160428,162417,1984,2023,2026,2027,217,2347
 5,2597,26330,2645,2821,29925,29968,3014,3059,3098,3099,3101,387712,471
 ,474382,5230,5232,55506,56648,6014,6122,6123,6124,6138,6152,6156,6187,
 6188,6207,6224,6233,6235,6470,6472,7311,80201,80755,83740,92815,9555
 GO:0034641

cellular amino acid biosynthetic process 0.0004087855668939752
 162417,216,29968,586,587,6470,6472 GO:0008652

biosynthetic process 0.0004138340532060459
 10399,10840,143244,158,160428,162417,1984,2023,2026,216,217,220,2
 3475,2597,2645,2821,29925,29968,3059,471,51102,5230,5232,55506,5564,55
 65,56648,5836,586,587,6014,6122,6123,6124,6138,6152,6156,6187,6188,620
 7,6224,6233,6235,6470,6472,7311,80755,8854,9555 GO:0009058

cellular macromolecule biosynthetic process 0.0005478001818329444
 10399,143244,1984,2597,2645,29925,56648,6122,6123,6124,6138,6152,
 6156,6187,6188,6207,6224,6233,6235,6470,6472,7311,80755 GO:0034645

carbohydrate phosphorylation 0.0010604272364423462
2645,3098,3099,3101,80201 GO:0046835

folic acid-containing compound metabolic process 0.0019927197908368374
10840,160428,471,6470,6472 GO:0006760

organic substance biosynthetic process 0.003163693453085785
10399,143244,158,162417,1984,2023,2026,216,217,220,23475,2597,264
5,2821,29925,29968,3059,471,51102,5230,5232,55506,5564,5565,56648,5836
,586,587,6014,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,6233,6
235,6470,6472,7311,80755,8854,9555 GO:1901576

pteridine-containing compound metabolic process 0.0048829062867107855
10840,160428,471,6470,6472 GO:0042558

positive regulation of translational elongation 0.005659593864818715
143244,1984,56648 GO:0045901

positive regulation of translational termination 0.005659593864818715
143244,1984,56648 GO:0045905

cellular metabolic process 0.006059870847800744
10399,10840,143244,158,160428,162417,1984,2017,2023,2026,2027,216
,217,220,23475,2597,26330,2645,2821,29925,29968,3014,3059,3098,3099,31
01,387712,4191,471,474382,51102,5230,5232,55506,5564,5565,56648,5834,5
836,586,587,6014,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,623
3,6235,6470,6472,7311,7314,7316,80201,80755,83740,8854,92815,9555
GO:0044237

alpha-amino acid biosynthetic process 0.006269160432328063
162417,29968,586,587,6470,6472 GO:1901607

primary metabolic process 0.008564453989666678
10399,143244,158,160428,162417,1984,2023,2026,2027,216,217,219,22
0,23475,2597,26330,2645,2821,29925,29968,3014,3059,3098,3099,3101,3877
12,4191,471,474382,51102,5230,5232,55506,5564,5565,56648,5834,5836,586
,587,6014,6122,6123,6124,6138,6152,6156,6187,6188,6207,6224,6233,6235,
6470,6472,7311,7314,7316,80201,80755,83740,8854,92815,9555 GO:0044238

glucose homeostasis 0.0312766633171484
10399,2645,2821,3098,3099,3101,5836,80201 GO:0042593

carbohydrate homeostasis 0.03255428199641142
10399,2645,2821,3098,3099,3101,5836,80201 GO:0033500

protein-containing complex organization 0.038178871538140736
10399,143244,1627,1984,2017,220,28988,3014,3059,474382,55506,5664
8,6122,6123,6152,6188,6470,6472,83740,8854,92815,9555 GO:0043933

organic substance metabolic process 0.0469424200675754
10399,10840,143244,158,160428,162417,1984,2023,2026,2027,216,217,
219,220,23475,2597,26330,2645,2821,29925,29968,3014,3059,3098,3099,310
1,387712,4191,471,474382,51102,5230,5232,55506,5564,5565,56648,5834,58

36, 586, 587, 6014, 6122, 6123, 6124, 6138, 6152, 6156, 6187, 6188, 6207, 6224, 6233
, 6235, 6470, 6472, 7311, 7314, 7316, 80201, 80755, 83740, 8854, 92815, 9555
GO:0071704

GO Enrichment Biological process for identified Human homologs of *rrp4-G226D*, *rrp40-W195R*, *rrp46-L191H* UP targets

synaptic vesicle priming 0.00001722747333646333 116841, 6616, 8773
GO:0016082

synaptic vesicle fusion to presynaptic active zone membrane
0.00002045632914476374 116841, 6616, 8773 GO:0031629

vesicle fusion to plasma membrane 0.000024064745532326275
116841, 6616, 8773 GO:0099500

synaptic vesicle membrane organization 0.00004272277217608498
116841, 6616, 8773 GO:0048499

exocytic process 0.0013127555557300235 116841, 6616, 8773 GO:0140029

vesicle fusion 0.0036237997981358318 116841, 6616, 8773 GO:0006906

synaptic vesicle exocytosis 0.003956170529847628 116841, 6616, 8773
GO:0016079

organelle membrane fusion 0.004071140885709456 116841, 6616, 8773
GO:0090174

synaptic vesicle cycle 0.005348362007655742 116841, 6616, 8773
GO:0099504

organelle fusion 0.009042138614174671 116841, 6616, 8773 GO:0048284

neurotransmitter secretion 0.009648660591108261 116841, 6616, 8773
GO:0007269

signal release from synapse 0.009648660591108261 116841, 6616, 8773
GO:0099643

membrane fusion 0.014132670537666155 116841, 6616, 8773 GO:0061025

vesicle-mediated transport in synapse 0.021194519283448055
116841, 6616, 8773 GO:0099003

neurotransmitter transport 0.028123807922711384 116841, 6616, 8773
GO:0006836

regulation of neurotransmitter levels 0.03028438204894036
116841, 6616, 8773 GO:0001505

regulated exocytosis 0.043327909978951584 116841, 6616, 8773 GO:0045055

4.8 References

1. Corbett, A.H., *Post-transcriptional regulation of gene expression and human disease*. Curr Opin Cell Biol, 2018. **52**: p. 96-104.
2. Makino, D.L., M. Baumgaertner, and E. Conti, *Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex*. Nature, 2013. **495**(7439): p. 70-75.
3. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'->5' exoribonucleases*. Cell, 1997. **91**(4): p. 457-466.
4. Mitchell, P., E. Petfalski, and D. Tollervey, *The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism*. Genes Dev, 1996. **10**(4): p. 502-13.
5. Bonneau, F., et al., *The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation*. Cell, 2009. **139**(3): p. 547-59.
6. Liu, Q., J.C. Greimann, and C.D. Lima, *Reconstitution, activities, and structure of the eukaryotic RNA exosome*. Cell, 2006. **127**(6): p. 1223-1237.
7. Wasmuth, E.V., K. Januszyk, and C.D. Lima, *Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA*. Nature, 2014. **511**(7510): p. 435-9.
8. Zinder, J.C., E.V. Wasmuth, and C.D. Lima, *Nuclear RNA Exosome at 3.1 A Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3*. Mol Cell, 2016. **64**(4): p. 734-745.
9. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 227-239.
10. Morton, D.J., et al., *The RNA exosome and RNA exosome-linked disease*. Rna, 2018. **24**(2): p. 127-142.
11. Schneider, C. and D. Tollervey, *Threading the barrel of the RNA exosome*. Trends in Biochemical Sciences, 2013. **38**(10): p. 485-493.
12. Chekanova, J.A., et al., *Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome*. Cell, 2007. **131**(7): p. 1340-53.
13. Gudipati, R.K., et al., *Extensive degradation of RNA precursors by the exosome in wild-type cells*. Mol Cell, 2012. **48**(3): p. 409-21.
14. Schneider, C., et al., *Transcriptome-wide Analysis of Exosome Targets*. Molecular Cell, 2012. **48**(3): p. 422-433.
15. Pefanis, E., et al., *Noncoding RNA transcription targets AID to divergently transcribed loci in B cells*. Nature, 2014. **514**(7522): p. 389-93.
16. Allmang, C., et al., *Functions of the exosome in rRNA, snoRNA and snRNA synthesis*. Embo Journal, 1999. **18**(19): p. 5399-5410.

17. Parker, R., *RNA Degradation in Saccharomyces cerevisiae*. Genetics, 2012. **191**(3): p. 671-702.
18. Wyers, F., et al., *Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase*. Cell, 2005. **121**(5): p. 725-737.
19. Kiss, D.L. and E.D. Andrulis, *Genome-wide analysis reveals distinct substrate specificities of Rrp6, Dis3, and core exosome subunits*. Rna, 2010. **16**(4): p. 781-791.
20. Belair, C., S. Sim, and S.L. Wolin, *Noncoding RNA Surveillance: The Ends Justify the Means*. Chemical Reviews, 2018. **118**(8): p. 4422-4447.
21. Molleston, J.M., et al., *A conserved virus-induced cytoplasmic TRAMP-like complex recruits the exosome to target viral RNA for degradation*. Genes & Development, 2016. **30**(14): p. 1658-1670.
22. Schuch, B., et al., *The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase*. Embo Journal, 2014. **33**(23): p. 2829-2846.
23. Weick, E.M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. Cell, 2018. **173**(7): p. 1663-1677 e21.
24. Stuparevic, I., et al., *Cotranscriptional Recruitment of RNA Exosome Cofactors Rrp47p and Mpp6p and Two Distinct Trf-Air-Mtr4 Polyadenylation (TRAMP) Complexes Assists the Exonuclease Rrp6p in the Targeting and Degradation of an Aberrant Messenger Ribonucleoprotein Particle (mRNP) in Yeast*. Journal of Biological Chemistry, 2013. **288**(44): p. 31816-31829.
25. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. Cell Reports, 2017. **20**(10): p. 2279-2286.
26. Milligan, L., et al., *A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts*. Molecular and Cellular Biology, 2008. **28**(17): p. 5446-5457.
27. Schilders, G., et al., *MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation*. Nucleic Acids Research, 2005. **33**(21): p. 6795-6804.
28. Wasmuth, E.V., et al., *Structure and reconstitution of yeast Mpp6-nuclear exosome complexes reveals that Mpp6 stimulates RNA decay and recruits the Mtr4 helicase*. Elife, 2017. **6**: p. 24.
29. LaCava, J., et al., *RNA degradation by the exosome is promoted by a nuclear polyadenylation complex*. Cell, 2005. **121**(5): p. 713-724.
30. Kadaba, S., et al., *Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae*. Genes Dev, 2004. **18**(11): p. 1227-40.
31. Milligan, L., et al., *A nuclear surveillance pathway for mRNAs with defective polyadenylation*. Mol Cell Biol, 2005. **25**(22): p. 9996-10004.
32. Lubas, M., et al., *Interaction profiling identifies the human nuclear exosome targeting complex*. Mol Cell, 2011. **43**(4): p. 624-37.
33. van Hoof, A., et al., *Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA*. Mol Cell Biol, 2000. **20**(21): p. 8230-43.

34. Wang, L., M.S. Lewis, and A.W. Johnson, *Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p*. RNA, 2005. **11**(8): p. 1291-302.
35. Liu, J.J., et al., *CryoEM structure of yeast cytoplasmic exosome complex*. Cell Res, 2016. **26**(7): p. 822-37.
36. Hou, D., M. Ruiz, and E.D. Andrulis, *The ribonuclease Dis3 is an essential regulator of the developmental transcriptome*. BMC Genomics, 2012. **13**.
37. Lim, S.J., et al., *Genome-wide localization of exosome components to active promoters and chromatin insulators in Drosophila*. Nucleic Acids Res, 2013. **41**(5): p. 2963-80.
38. Lorentzen, E., et al., *RNA channelling by the archaeal exosome*. Embo Reports, 2007. **8**(5): p. 470-476.
39. Fasken, M.B., et al., *The RNA Exosome and Human Disease*. Methods Mol Biol, 2020. **2062**: p. 3-33.
40. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. J Neurol, 2013. **260**(7): p. 1866-70.
41. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. Nature Communications, 2014. **5**.
42. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
43. Di Donato, N., et al., *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*. Journal of Medical Genetics, 2016. **53**(6): p. 419-425.
44. Eggens, V.R.C., et al., *EXOSC3 mutations in pontocerebellar hypoplasia type I: novel mutations and genotype-phenotype correlations*. Orphanet Journal of Rare Diseases, 2014. **9**(1): p. 23.
45. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
46. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nature Genetics, 2012. **44**(6): p. 704-U134.
47. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. Hum Mol Genet, 2020. **29**(13): p. 2218-2239.
48. Somashekar, P.H., et al., *Bi-allelic missense variant, p.Ser35Leu in EXOSC1 is associated with pontocerebellar hypoplasia*. Clin Genet, 2021. **99**(4): p. 594-600.
49. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nat Genet, 2012. **44**(6): p. 704-8.
50. Bizzari, S., et al., *Expanded PCHID phenotype linked to EXOSC9 mutation*. Eur J Med Genet, 2020. **63**(1): p. 103622.

51. Burns, D., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. *Neuromuscul Disord*, 2017. **27**: p. S38.
52. Rudnik-Schoneborn, S., et al., *Pontocerebellar hypoplasia type 1: clinical spectrum and relevance of EXOSC3 mutations*. *Neurology*, 2013. **80**(5): p. 438-46.
53. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. *Nature genetics*, 2012. **44**(6): p. 704-8.
54. Sterrett, M.C., et al., *A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex*. *Rna*, 2021. **27**(9): p. 1046-1067.
55. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. *Genetics*, 2017. **205**(1): p. 221-+.
56. Gillespie, A., et al., *Mutations of EXOSC3/Rp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in S. cerevisiae*. *RNA*, 2017. **23**(4): p. 466-472.
57. Morton, D.J., B., Jalloh, I., Kremsky, L., Kim, T., Le, K., Nguyen, J.C., Rounds, M.C., Sterrett, B., Brown, S.W., Leung, M.B., Fasken, K.H., Moberg, A.H., Corbett, *A Drosophila Model of Pontocerebellar Hypoplasia Reveals a Critical Role for the RNA Exosome in Neurons*. *BioRxiv* 727693 2019. **[Preprint]**
58. Mitchell, P., et al., *The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases*. *Cell*, 1997. **91**(4): p. 457-66.
59. Belair, C., et al., *The RNA exosome nuclease complex regulates human embryonic stem cell differentiation*. *J Cell Biol*, 2019.
60. Adams, A., et al., *Methods in Yeast Genetics*. 1997, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
61. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*. Second ed. 1989, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
62. Schaeffer, D., et al., *The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities*. *Nat Struct Mol Biol*, 2009. **16**(1): p. 56-62.
63. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae*. *Genetics*, 1989. **122**(1): p. 19-27.
64. Da, B., D. Dawson, and T. Stearns, *Methods In Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. 2000.
65. Bogdanova, E.A., et al., *Normalization of Full-Length-Enriched cDNA, in cDNA Libraries: Methods and Applications*, C. Lu, J. Browse, and J.G. Wallis, Editors. 2011, Humana Press: Totowa, NJ. p. 85-98.
66. Pertea, M., et al., *StringTie enables improved reconstruction of a transcriptome from RNA-seq reads*. *Nature Biotechnology*, 2015. **33**(3): p. 290-295.

67. Kolde, R., *Pheatmap: pretty heatmaps*. R package version, 2012. **1**(2): p. 726.
68. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features*. *Bioinformatics*, 2013. **30**(7): p. 923-930.
69. Engel, S.R., et al., *The reference genome sequence of Saccharomyces cerevisiae: then and now*. G3 (Bethesda), 2014. **4**(3): p. 389-98.
70. Xu, Z., et al., *Bidirectional promoters generate pervasive transcription in yeast*. *Nature*, 2009. **457**(7232): p. 1033-7.
71. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biology*, 2014. **15**(12): p. 550.
72. Zhu, A., J.G. Ibrahim, and M.I. Love, *Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences*. *Bioinformatics*, 2019. **35**(12): p. 2084-2092.
73. Blighe, K., S. Rana, and M. Lewis, *EnhancedVolcano: Publication-Ready Volcano Plots with Enhanced Colouring and Labeling*. 2018 Available online: <https://github.com/kevinblighe>. EnhancedVolcano (accessed on 14 December 2020).
74. Conway, J.R., A. Lex, and N. Gehlenborg, *UpSetR: an R package for the visualization of intersecting sets and their properties*. *Bioinformatics*, 2017. **33**(18): p. 2938-2940.
75. Cherry, J.M., et al., *Saccharomyces Genome Database: the genomics resource of budding yeast*. *Nucleic Acids Res*, 2012. **40**(Database issue): p. D700-5.
76. Briggs, M.W., K.T. Burkard, and J.S. Butler, *Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation*. *J Biol Chem*, 1998. **273**(21): p. 13255-63.
77. Marin-Vicente, C., et al., *RRP6/EXOSC10 is required for the repair of DNA double-strand breaks by homologous recombination*. *Journal of Cell Science*, 2015. **128**(6): p. 1097-1107.
78. Bush, G.L. and D.I. Meyer, *The refolding activity of the yeast heat shock proteins Ssa1 and Ssa2 defines their role in protein translocation*. *J Cell Biol*, 1996. **135**(5): p. 1229-37.
79. Unno, K., et al., *Role of Hsp70 subfamily, Ssa, in protein folding in yeast cells, seen in luciferase-transformed ssa mutants*. *Biol Pharm Bull*, 1997. **20**(12): p. 1240-4.
80. López-Ribot, J.L. and W.L. Chaffin, *Members of the Hsp70 family of proteins in the cell wall of Saccharomyces cerevisiae*. *J Bacteriol*, 1996. **178**(15): p. 4724-6.
81. Brown, C.R., J.A. McCann, and H.L. Chiang, *The heat shock protein Ssa2p is required for import of fructose-1, 6-bisphosphatase into Vid vesicles*. *J Cell Biol*, 2000. **150**(1): p. 65-76.
82. Borkovich, K.A., et al., *hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures*. *Mol Cell Biol*, 1989. **9**(9): p. 3919-30.
83. Gross, D.S., et al., *Promoter function and in situ protein/DNA interactions upstream of the yeast HSP90 heat shock genes*. *Antonie Van Leeuwenhoek*, 1990. **58**(3): p. 175-86.

84. Planta, R.J. and W.H. Mager, *The list of cytoplasmic ribosomal proteins of Saccharomyces cerevisiae*. Yeast, 1998. **14**(5): p. 471-7.
85. Mrsa, V. and W. Tanner, *Role of NaOH-extractable cell wall proteins Ccw5p, Ccw6p, Ccw7p and Ccw8p (members of the Pir protein family) in stability of the Saccharomyces cerevisiae cell wall*. Yeast, 1999. **15**(10a): p. 813-20.
86. Kobayashi, N., et al., *Structure and functional analysis of the multistress response gene DDR2 from Saccharomyces cerevisiae*. Biochem Biophys Res Commun, 1996. **229**(2): p. 540-7.
87. Lacroute, F., *Regulation of pyrimidine biosynthesis in Saccharomyces cerevisiae*. J Bacteriol, 1968. **95**(3): p. 824-32.
88. Minard, K.I. and L. McAlister-Henn, *Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from Saccharomyces cerevisiae: evidence for three isozymes of yeast malate dehydrogenase*. Mol Cell Biol, 1991. **11**(1): p. 370-80.
89. Keesey, J.K., Jr., R. Bigelis, and G.R. Fink, *The product of the his4 gene cluster in Saccharomyces cerevisiae. A trifunctional polypeptide*. J Biol Chem, 1979. **254**(15): p. 7427-33.
90. Yamamoto, R.T., et al., *RRN3 gene of Saccharomyces cerevisiae encodes an essential RNA polymerase I transcription factor which interacts with the polymerase independently of DNA template*. Embo j, 1996. **15**(15): p. 3964-73.
91. Hall, C. and F.S. Dietrich, *The reacquisition of biotin prototrophy in Saccharomyces cerevisiae involved horizontal gene transfer, gene duplication and gene clustering*. Genetics, 2007. **177**(4): p. 2293-307.
92. Phalip, V., et al., *Characterization of the biotin biosynthesis pathway in Saccharomyces cerevisiae and evidence for a cluster containing BIO5, a novel gene involved in vitamer uptake*. Gene, 1999. **232**(1): p. 43-51.
93. Mörtl, S., et al., *Biosynthesis of riboflavin. Lumazine synthase of Escherichia coli*. J Biol Chem, 1996. **271**(52): p. 33201-7.
94. Rieger, K.J., et al., *A novel nuclear gene, CBT1, essential for mitochondrial cytochrome b formation: terminal processing of mRNA and intron dependence*. Current Genetics, 1997. **32**(3): p. 163-174.
95. Ellis, T.P., M.S. Schonauer, and C.L. Dieckmann, *CBT1 interacts genetically with CBP1 and the mitochondrially encoded cytochrome b gene and is required to stabilize the mature cytochrome b mRNA of Saccharomyces cerevisiae*. Genetics, 2005. **171**(3): p. 949-57.
96. Wang, C., et al., *Rrp6 Moonlights in an RNA Exosome-Independent Manner to Promote Cell Survival and Gene Expression during Stress*. Cell Reports, 2020. **31**(10): p. 107754.
97. Gasch, A.P., et al., *Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p*. Mol Biol Cell, 2001. **12**(10): p. 2987-3003.
98. Ribeiro, R.A., N. Bourbon-Melo, and I. Sá-Correia, *The cell wall and the response and tolerance to stresses of biotechnological relevance in yeasts*. Front Microbiol, 2022. **13**: p. 953479.

99. Davis, C.A. and M. Ares, Jr., *Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3262-7.
100. Gudipati, R.K., et al., *Extensive Degradation of RNA Precursors by the Exosome in Wild-Type Cells*. Molecular Cell, 2012. **48**(3): p. 409-421.
101. Yamada, M., et al., *Y'-Help1, a DNA helicase encoded by the yeast subtelomeric Y' element, is induced in survivors defective for telomerase*. J Biol Chem, 1998. **273**(50): p. 33360-6.
102. Coy, S., et al., *The Sm complex is required for the processing of non-coding RNAs by the exosome*. PloS one, 2013. **8**(6): p. e65606-e65606.
103. Schmid, M. and T.H. Jensen, *The Nuclear RNA Exosome and Its Cofactors*. Adv Exp Med Biol, 2019. **1203**: p. 113-132.
104. Ben-Shem, A., et al., *The structure of the eukaryotic ribosome at 3.0 Å resolution*. Science, 2011. **334**(6062): p. 1524-9.
105. Fox, M.J., et al., *The Exosome Component Rrp6 Is Required for RNA Polymerase II Termination at Specific Targets of the Nrd1-Nab3 Pathway*. PLOS Genetics, 2015. **11**(2): p. e1004999.
106. Wyers, F., et al., *Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase*. Cell, 2005. **121**(5): p. 725-737.
107. Arigo, J.T., et al., *Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3*. Mol Cell, 2006. **23**(6): p. 841-51.
108. Schulz, D., et al., *Transcriptome surveillance by selective termination of noncoding RNA synthesis*. Cell, 2013. **155**(5): p. 1075-87.
109. Thiebaut, M., et al., *Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance*. Mol Cell, 2006. **23**(6): p. 853-64.
110. Halevy, A., et al., *Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia*. J Neurol, 2014. **261**(11): p. 2165-9.
111. Schuller, J.M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. Science, 2018. **360**(6385): p. 219-222.
112. Oddone, A., et al., *Structural and biochemical characterization of the yeast exosome component Rrp40*. EMBO Rep, 2007. **8**(1): p. 63-9.
113. Di Donato, N., et al., *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, 2016. **53**(6): p. 419-25.

Chapter V: Discussion & Future Directions

5.1 Summary of Presented Studies

Recent clinical studies linking mutations in genes encoding structural subunits of the RNA exosome to distinct, tissue-specific diseases have presented an interesting, biological question regarding the function of the complex *in vivo* [1]. A growing collection of identified RNA exosome-linked diseases, classified as “RNA exosomopathies”, range in clinical manifestations and do not appear to correlate with the type of structural gene that is mutated. What is common between these different RNA exosomopathies is that the causative mutations encode pathogenic amino acid substitutions in conserved domains of the structural subunits of the RNA exosome. The distinct patient phenotypes of each RNA exosomopathy suggest specific molecular consequences resulting from each of these amino acid substitutions. This raises the question as to how single amino acid substitutions in highly conserved, structural subunits of the RNA exosome differentially affect the function of this essential complex, and, consequently, underlie distinct disease pathologies.

Using the *Saccharomyces cerevisiae* genetic model system, I modeled patient RNA exosomopathy mutations in orthologous genes and assessed the molecular and functional defects resulting from different pathogenic amino acid substitutions. I hypothesize that disease-linked amino acid substitutions within structural subunits could be altering RNA exosome function differentially through three potential yet nonexclusive mechanisms (summarized by Chapter I Figure 4). These disease-linked amino acid substitutions could differentially affect (I) the integrity of the complex, (II) key interactions with cofactor proteins, and/or (III) directly disrupt processing/degradation of RNAs by this molecular machine. Any of these three mechanisms would result in compromised processing and/or degradation of RNA exosome targets and, ultimately, lead to transcriptomic changes. Utilizing the budding yeast platform, I took genetic and

biochemical approaches to assess these mechanisms and found that RNA exosome function is impacted differentially by pathogenic amino acid substitutions *in vivo*.

In Chapter II, I modeled and analyzed pathogenic amino acid substitutions in the *S. cerevisiae* EXOSC2 ortholog, Rrp4, generating *rrp4-G58V* and *rrp4-G226D* mutants. These mutants correspond to the SHRF-linked mutations *EXOSC2-G30V* and *EXOSC2-G198D*, respectively. My work demonstrates that the *rrp4-G58V* mutation is lethal in *S. cerevisiae*, while the *rrp4-G226D* mutation results in severe growth defects and increases in steady-state levels of known direct RNA exosome targets, suggesting RNA exosome function is impaired. Genetic studies suggest interactions with key RNA exosome cofactors are impaired in these *rrp4-G226D* cells, particularly within context of interactions with the essential RNA helicase Mtr4. Mtr4 is a key nuclear RNA exosome cofactor that, among other roles, is necessary for RNA exosome processing of ribosomal RNA precursors [2-5]. RNA-Seq analysis of the *rrp4-G226D* cells reveals large transcriptomic changes resulting from this modeled RNA exosomopathy mutation, further suggesting impaired function of the RNA exosome complex. Biochemical analyses revealed that the Rrp4 G226D variant could associate within the RNA exosome complex, however, this mutant variant could not efficiently compete with wild-type Rrp4 for incorporation into the RNA exosome complex. Furthermore, these data reveal a decreased interaction between the Rrp4 G226D variant and the essential RNA helicase Mtr4, supporting the genetic experiments in suggesting this pathogenic amino acid substitution causes a decreased association of the RNA exosome complex with Mtr4 *in vivo*. From these data, I can conclude that the Rrp4 G226D variant likely impacts key cofactor interactions and may also impact overall complex integrity.

In Chapter III, I modeled and analyzed a multiple myeloma patient *EXOSC2* mutation in the *S. cerevisiae* homolog *RRP4*, generating *rrp4-M68T* mutant cells expressing the variant Rrp4

M68T. The *rrp4-M68T* cells show growth defects when grown in media containing drugs that impact RNA processing and the *rrp4-M68T* cells have accumulation of known RNA exosome targets, suggesting the multiple myeloma associated amino acid substitution does affect function of the complex. Taking similar approaches as those performed in Chapter II, I used genetic and biochemical studies to assess whether the *rrp4-M68T* mutation impacts interactions with known RNA exosome cofactors. The *rrp4-M68T* cells show distinct negative genetic interactions with *mpp6* and *mtr4* mutants, both nuclear RNA exosome cofactors [3, 6-8]. Biochemical assays reveal that in fact the M68T substitution in Rrp4 decreases association with the Mtr4 helicase. This finding led me to conclude that the defects in RNA exosome function observed within the *rrp4-M68T* cells could result from an impaired interaction between the complex and the essential RNA helicase Mtr4. Furthermore, these data suggest that the introduction of the multiple myeloma associated amino acid change could impact the binding interface between EXOSC2 and MTR4, potentially impairing the function of the essential RNA exosome *in vivo* for a subset of Mtr4-dependent targets.

In Chapter IV, I performed an unbiased RNA-Seq experiment across a collection of *rrp* mutant models to extensively compare different RNA exosomopathy mutations *in vivo*. From these data, I determined that the *rrp4-G226D*, the *rrp40-W195R* and the *rrp46-L191H* mutant models have the most significant consequences in yeast, with all three showing a large amount of differentially expressed genes as identified through differential expression analysis. The *rrp4-G226D* and *rrp40-W195R* cells model mutations in *EXOSC2* and *EXOSC3* that are linked to SHRF and PCH [9, 10], respectively. The *rrp46-L191H* cells model an *EXOSC5* mutation that causes cerebellar atrophy and neurological defect [11]. The RNA-Seq experiment provides many

intriguing results but one of interest is the implication that rRNA processing/modifications and ribosome biogenesis are potential pathways impacted within all three *rrp* mutant cells.

5.2 Conclusions from Presented Studies

Taken together, these studies demonstrate that different disease-linked amino acid substitutions in the structural components of the RNA exosome result in functional consequences *in vivo*. By comparing between *rrp* mutant models, I am also able to conclude that some of these *in vivo* consequences are shared across the *rrp* mutants while some also appear distinct. These comparative analyses also provide further insight into the biology that may underlie the different disease pathologies linked to RNA exosomopathy mutations. In Chapter II, I compared the functional and molecular consequences cells resulting from the the *rrp4-G226D* mutation to those previously observed in a different *rrp* mutant that models a PCH-linked *EXOSC3* mutation, *rrp40-W195R* [12]. Both exosomopathy mutant models show growth defects and changes in steady state levels of RNA transcripts that are targeted by the RNA exosome however some transcripts show statistically significant changes only in *rrp4-G226D* cells. Genetic analyses also suggests that *rrp4-G226D* and the *rrp40-W195R* cells have similar negative genetic interactions with key nuclear cofactors mutants, however there are distinct negative genetic interactions between *mtr4* mutants in *rrp4-G226D* cells that are not observed within *rrp40-W195R* cells. This comparative analysis suggests that the missense mutations in *RRP4* and *RRP40* have distinct functional consequences for RNA exosome activity, which would be consistent with the distinct clinical presentations in patients with these pathogenic variants. However, these data also suggested that the *rrp4-G226D* allele may simply be a stronger allele than *rrp40-W195R*. The comparative RNA-Seq analysis presented in Chapter IV provides evidence that there are some differences in the functional consequences between these modeled cap RNA exosomopathy mutations. I do see

larger transcriptomic changes in *rrp4-G226D* cells compared to the *rrp40-W195R* cells. Additionally, some of those transcriptomic changes are shared and linked to translation and telomere maintenance. However, I also observe many transcripts that are changed solely in the *rrp4-G226D* or *rrp40-W195R* cells. Gene Ontology analyses reveal that those targets unique to the *rrp40-W195R* cells are enriched in processes involved in translation and ribosome biogenesis while targets unique to the *rrp4-G226D* cells are enriched more in metabolic and biosynthetic processes. Integrating the results presented in both Chapter II and Chapter IV suggest that the *rrp4-G226D* and *rrp40-W195R* mutations result in shared and unique consequences in the function of the RNA exosome. Further exploration of these unique transcripts impacted in the *rrp4-G226D* and *rrp40-W195R* cells may shed light on differences that could underlie the different disease pathologies observed in SHRF and PCH patients.

In Chapter III, I compare the *in vivo* consequences resulting from the *rrp4-M68T* mutation to those of the *rrp4-G226D* mutation. I observe differences in severity of the functional and molecular consequences between these two *rrp4* mutant models, with *rrp4-G226D* cells having more severe defects in growth and having higher steady state levels of more RNA exosome target transcripts. I do observe increases in extended forms of noncoding RNAs within both *rrp4-M68T* and *rrp4-G226D* cells. However, I see that the *rrp4-M68T* mutant models do not share steady state level increases in select CUTs or 5.8S rRNA precursor as observed in *rrp4-G226D* cells. This leads me to predict that the two *rrp4* mutant models have some similar defects in RNA exosome processing due to destabilized Mtr4 interactions. However the *rrp4-G226D* mutant models have more severe consequences, perhaps due to the potential defects in complex integrity as was concluded from data in Chapter III. The difference in severity of functional and molecular consequences between the *rrp4-M68T* and *rrp4-G226D* mutant models may partially explain the

differences in disease pathology between SHRF patients and the multiple myeloma patient with the identified *EXOSC2* mutation. The SHRF linked *EXOSC2* mutations are classified as causative; In contrast, the *EXOCS2* mutation linked to multiple myeloma is a spontaneous, somatic mutation that likely co-occurred with chromosomal aberrations. This suggests the *EXOSC2* mutation that is modeled in the *rrp4-M68T* model is a passenger mutation rather than a pathogenic driver of the multiple myeloma.

The work presented in Chapter IV represents the first *in vivo* comparative study of a large collection of RNA exosomopathy mutant models. While these data do not assess all potential mechanisms by which the different pathogenic amino acid substitutions may impact the RNA exosome complex, these datasets allow me to assess whether the transcriptomic changes for each *rrp* mutant model are distinct and, thus, if there are differential broad impacts on the processing or degradation of RNA targets *in vivo*. The information captured in the RNA-Seq data opens the door to new directions of research that can further expand our understanding of the *in vivo* consequences resulting from these different RNA exosomopathy mutations and, thus, what processes may underlie the different disease pathologies.

5.3 Future Directions

In thinking about future directions, the link between modeled RNA exosomopathy mutations and impacted rRNA processing/ribosome presented in Chapter IV provides a very exciting, new research question. This research question is of particular interest when we compare the clinical manifestations of RNA exosomopathies to those of ribosomopathies, a diverse class of diseases linked to mutation in ribosomal components [13, 14]. The ribosome produces all cellular proteins and is a large complex containing numerous conserved components [15, 16]. The

eukaryotic ribosome consists of a small (40S) and large (60S) RNA-protein subunits. In *Saccharomyces cerevisiae*, the 40S subunit consists of 33 ribosomal proteins (Rps prefix) and the 18S ribosomal RNA (rRNA), while the 60S subunit (60S) consists of 46 ribosomal proteins (Rpl prefix) and 3 rRNAs (25S, 5.8S and 5S rRNA).

Similar to RNA exosomopathies, ribosomopathies are linked to mutations in genes that encode components of the small or large subunits. Ribosomopathies comprise a wide variety of diseases with distinct clinical phenotypes, including developmental abnormalities, hematopoietic defects, craniofacial malformations, short stature, mental and motor retardation, and molecular oncogenic hallmarks [13, 14]. Several ribosomopathies also have neurological pathologies such as PEHO (progressive encephalopathy with oedema, hypsarrhythmia, and optic atrophy) and Bowen–Conradi syndromes [17, 18]. The diversity in clinical manifestations of RNA exosomopathies mirrors that seen in ribosomopathies, with several RNA exosomopathy diseases and ribosomopathy disorders having prevalent neurological defects.

As the RNA exosome and the ribosome both support cellular processes essential for every living cell [19, 20], defects in their production or function are expected to have global lethal effects rather than cause tissue-specific pathologies. More evidence linking these disease classes would be the role the RNA exosome plays in processing rRNA, particularly the 5.8S rRNA, which is a major component of the small ribosome subunit [21, 22]. In fact, rRNA comprises a major portion of the ribosome, and thus any changes in the levels or processing of the nucleic acid would likely impact the composition and potentially the function of the molecular machine [16, 23, 24].

The clinical similarities, the conserved role the RNA exosome plays in processing rRNA, and the RNA-Seq results presented in Chapter IV, all suggest that RNA exosomopathies and ribosomopathies share common molecular mechanisms that impact translation and could

contribute to or underlie pathology. Therefore I predict that some molecular defects resulting from the RNA exosomopathy mutations *in vivo* may arise from translation dysregulation. As modeled in Figure 1, there could be several ways that changes in RNA exosome function resulting from RNA exosomopathy pathogenic amino acid substitutions could result in translational defects. One potential mechanism is a decrease of ribosomal protein genes or ribosomal assembly factors resulting from the RNA exosomopathy mutations that ultimately would result in altered ribosome biogenesis. As observed within the comparative RNA-Seq data in Chapter IV, there are broad transcriptomic changes within three of the *rrp* mutant models that include decreased ribosomal protein gene expression. We cannot conclude yet mechanistically how these modeled pathogenic amino acid substitutions cause decreased levels of mRNA transcripts that encode ribosomal proteins. However, previous studies have shown similar broad decreases in ribosomal protein gene expression in cells with disrupted RNA exosome function [25]. Regardless of how exactly the RNA exosome impacts levels of ribosomal protein gene mRNAs, a broad decrease as observed in our *rrp* mutant models would likely impact ribosomal biogenesis, resulting in perhaps fewer and/or defective ribosomes and, thus, consequence for translation.

Another potential mechanism as to how these RNA exosomopathy pathogenic amino acid substitutions result in translational dysregulation could be through impacting RNAs that are important for rRNA modifications, such as small nucleolar RNAs. In addition to its critical role in processing rRNA, the nuclear RNA exosome is a central factor in generating stable small nucleolar RNA species [26]. Recent studies have also suggested that the RNA exosome aids in the turnover of these RNA species particularly through interactions with known nuclear cofactors [26, 27]. Small nucleolar RNAs, or snoRNAs, play an essential role in post-transcriptional modification of rRNA [28, 29]. The snoRNA act as a guide for enzymes, resulting in modifications of specific

ribonucleotides within the rRNA [28, 29]. Post-transcriptional modifications of rRNA most commonly associated with snoRNA action include 2'-O-methylation and pseudo-uridylation [30], modifications which are thought to generally help stabilize the structure of the ribosome or facilitate efficient translation [31]. Interestingly, multiple snoRNA species are commonly increased across the *rrp* mutant models assessed in the comparative RNA-Seq experiment presented in Chapter IV. Furthermore, Gene Ontology (GO) enrichment analyses of differentially expressed genes in the *rrp* mutant cells repeatedly revealed rRNA modification and rRNA methylation as biological processes significantly impacted, particularly when analysis was performed on the transcripts that are increased in comparison the wild-type control cells. Therefore, it is possible that the different pathogenic amino acid substitutions are impacting the function of the RNA exosome in processing or degrading these snoRNAs, thus changing the modification patterns of rRNA. Changes in rRNA modifications could ultimately affect the integrity of the ribosome complex, thereby changing the translation efficiency of the molecular machines. Currently how these pathogenic amino acid substitutions impact snoRNA is not known. However, as the processing and degradation of snoRNAs by the RNA exosome is thought to be mediated through interactions with certain nuclear cofactors [26, 27], it is reasonable to hypothesize that perhaps these different pathogenic amino acid substitutions disrupt these interactions and consequently cause changes in snoRNAs within the *rrp* cells. Interestingly, “orphan” snoRNAs have been identified recently. Orphan snoRNAs are thought to play noncanonical functional roles on other nucleic acid molecules, including splicing and polyadenylation of pre-mRNA, and even stability of chromatin [28, 32, 33]. Several orphan snoRNAs have been identified in neuronal tissue, suggesting that the substrate repertoire and cellular influence of snoRNAs is much broader than previously appreciated [28]. As such, impacts

on snoRNAs by pathogenic amino acid substitutions in the RNA exosome may stretch beyond just impacting rRNA modifications in human patients, and may be of interest to investigate further given the neurological pathologies of many identified RNA exosomopathies.

A third potential mechanism as to how these RNA exosomopathy pathogenic amino acid substitutions result in translational dysregulation could be by directly impacting the processing of the rRNA. As mentioned, the RNA exosome plays a pivotal role in generating mature 5.8S rRNA [34, 35]. *In vivo* studies of *rrp* mutant models, such as those described in Chapter II, revealed that the modeled pathogenic amino acid substitutions result in accumulation of premature 5.8S rRNA [11, 12, 36]. As described previously, the 5.8S rRNA is one of three rRNAs, together with several ribosomal proteins, that comprise the large 60S subunit of the eukaryotic ribosome. Within the nucleolus, the 5.8S rRNA is transcribed as a single, large precursor rRNA that also contains 18S and 25S rRNA. This long precursor rRNA is cleaved and processed by various endo- and exonucleolytic processing reactions, including 3' end processing by the RNA exosome, to generate the mature 5.8S rRNA [37, 38]. Concurrent with pre-rRNA processing, pre-rRNA are modified and subjected to structural rearrangements, allowing for assembly with ribosomal proteins to build the ribosomal small and large subunits [37-40]. This multi-step, compartmentalized process of ribosome biogenesis has evolved to safeguard the cell from aberrantly assembled ribosomal subunits, which, if engaged in translation, could have deleterious consequences, such as sequestering translation factors, stalling elongation, and/or reducing fidelity of protein synthesis [41, 42]. A previous study that assessed functional mutants of the RNA exosome observed large ribosomal subunits containing premature 5.8S rRNA not only accumulate but engage with translation elongation machinery in the cytoplasm, thus generating a pool of “immature” ribosomes that were actively translating [43]. Given that our RNA exosomopathy mutant models

show accumulation of premature 5.8S rRNA species [11, 12, 36], these cells could have “immature” ribosomes that have the ability to engage in translation. Given the extensive body of work showing how aberrant ribosomes disrupt overall translation fidelity [41, 42], it stands to reason that the *rrp* mutant models could be having translational defects due to this pool of ribosomes containing premature or misprocessed rRNA.

Ultimately, a change in translation or translational fidelity, as depicted in Figure 1, could underline patient pathology. Thus, translational fidelity changes could be one of many consequences resulting from these pathogenic amino acid substitutions in RNA exosome structural subunits. Given the differential growth defects of *rrp* mutant cells presented in Chapter IV, and the diverse clinical pathologies of RNA exosomopathy patients [9-11, 44-54], some of these RNA exosomopathy mutations may impact several different pathways—some that may be shared and some that may be unique to the specific pathogenic amino acid substitution. Another consideration in understanding the unique disease pathologies of RNA exosomopathies is describing RNA exosome requirements during certain developmental timepoints. Most, if not all, RNA exosomopathies include disorders that include neurodevelopmental delays. Pontocerebellar hypoplasia (PCH), which presents in patients with mutations in *EXOSC1*, *EXOSC3*, *EXOSC8*, *EXOSC9*, generally has a prenatal onset [55], further suggesting that some molecular defects in RNA exosomopathy patients are developmental. Recent studies have also begun linking RNA exosome function to cellular differentiation, further suggesting that the complex has differential impacts based on the plasticity of the cell [56-58]. Therefore, some pathogenic amino acid substitutions may have molecular and functional consequences that are most pronounced when the cell has specific requirements for the RNA exosome, such as those that could exist during key

developmental time points. This could then result in the differential impacts on tissues and/or the diversity of disease pathologies seen in RNA exosomopathy patients.

Lastly, there are still several mechanistic questions that exist with these RNA exosomopathy mutations. In particular, assessing how the pathogenic amino acid substitutions impact the integrity of the complex is an active area of research. In my work, I performed some biochemical assays that generally assessed whether mutant Rrp subunit variants harboring the disease-linked pathogenic amino acid substitutions had different protein steady-state levels from the wild-type Rrp subunit and whether the mutant Rrp variant could associate with the complex. While informative, the complex is not static and may have different dynamics depending on the presence of the pathogenic amino acid substitution. Therefore, a more in depth approach to assess complex integrity defects would be useful to determine whether subunit association/disassociation with the complex is impacted by the presence of a pathogenic amino acid substitution. Preliminary data available in Appendix II tests an experimental design that measures exchange between the mutant Rrp variant and a wild-type Rrp control over time, utilizing our *rrp* mutant models and an inducible construct. An experiment similar would not only provide insight into what *in vivo* molecular consequences occur due to the pathogenic amino acid substitution but also shed some light on the biochemistry of the complex.

5.4 Closing Remarks

In its entirety, the work presented here begins elucidating the functional consequences that result from disease-associated mutations identified in genes encoding the structural subunits of the RNA exosome. Furthermore, the comparative data analysis I could pursue utilizing *Saccharomyces cerevisiae* can serve as a foundation to systematically investigate each exosomopathy mutation and directly compare their impacts on the biology of the complex. My

work adds to a robust platform generated within the Corbett lab to understand how pathogenic amino acid substitutions impact the function of the RNA exosome *in vivo*. Additionally, my work represents a commitment to mentorship and training, as several of these experiments were done in tandem with undergraduate mentees. The hands-on research opportunities provided through *S. cerevisiae* studies can have broad impacts on developing a diverse STEM field as it gives trainees autonomy and, thus, confidence in continuing in academic research.

Continuing these types of *in vivo* functional studies that use a powerful genetic system such as *S. cerevisiae* will also shed light on the basic biology of the RNA exosome, as many questions still exist regarding RNA substrate targeting and regulation of the essential molecular machine. Moreover, utilizing *S. cerevisiae* allows for iterative and expandable comparative analyses as new RNA exosomopathy mutations are identified regularly. These *in vivo* functional analyses and comparative approaches will allow for deeper understanding of how diverse clinical symptoms are linked to changes in the singular, conserved RNA exosome molecular machine.

5.5 Chapter V Figures

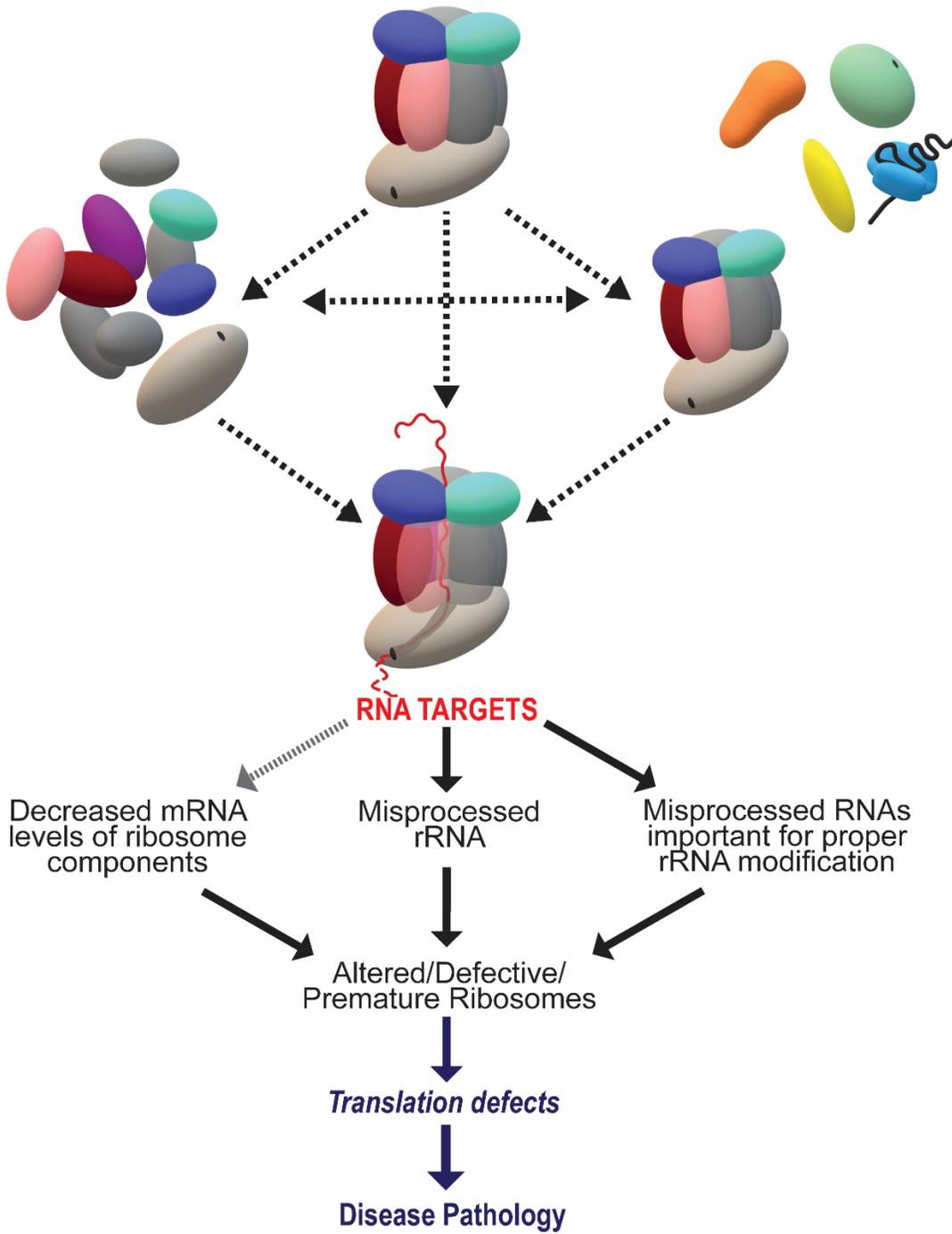


Figure 1. Proposed model of how RNA exosomopathy amino acid substitutions may result in distinct translational defects that could underlie patient pathology.

Disease-linked amino acid substitutions may differentially impact translation through several mechanisms. One mechanism is by decreasing levels of mRNAs that encode ribosomal proteins. Evidence for RNA exosome dysfunction resulting in decreased ribosomal protein genes is provided by the comparative RNA-Seq analysis presented in Chapter IV, in which I observe shared decreases in *RPS* and *RPL* transcripts across many *rrp* mutant cells. Previous studies have also shown similar broad decreases in ribosomal protein gene expression in cells with disrupted RNA exosome function [25]. However the exact mechanism of how defects in RNA exosome function result in decreased ribosomal protein mRNA levels remains unknown (represented as gray dotted line). More direct mechanisms that would explain defects in RNA exosome function resulting from pathogenic amino acid substitutions and translation changes are misprocessing ribosomal RNA (rRNA) and/or small nucleolar RNA (snoRNA). Both rRNA, particularly precursors to the 5.8S rRNA, and snoRNAs are known targets of the RNA exosome. Coincidentally, our RNA exosomopathy mutant models show accumulation of pre-5.8S rRNA (as shown in Chapter II and in previous studies [11, 12]) and increased steady state levels of snoRNAs (as shown in Chapter IV). All three potential mechanisms would ultimately lead to altered, defective or immature ribosomes which could have downstream consequences in translational fidelity *in vivo*. Changes in translation could underlie some of the diverse pathologies observed in RNA exosomopathy patients.

5.6 References

1. Fasken, M.B., et al., *The RNA Exosome and Human Disease*. Methods Mol Biol, 2020. **2062**: p. 3-33.
2. Belair, C., S. Sim, and S.L. Wolin, *Noncoding RNA Surveillance: The Ends Justify the Means*. Chem Rev, 2018. **118**(8): p. 4422-4447.
3. de la Cruz, J., et al., *Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae*. Embo j, 1998. **17**(4): p. 1128-40.
4. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. Cell Reports, 2017. **20**(10): p. 2279-2286.
5. Jackson, R.N., et al., *The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing*. EMBO J, 2010. **29**(13): p. 2205-16.
6. Milligan, L., et al., *A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts*. Mol Cell Biol, 2008. **28**(17): p. 5446-57.
7. Schuch, B., et al., *The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase*. Embo Journal, 2014. **33**(23): p. 2829-2846.
8. Weick, E.M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. Cell, 2018. **173**(7): p. 1663-1677 e21.
9. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nature Genetics, 2012. **44**(6): p. 704-U134.
10. Di Donato, N., et al., *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*. Journal of Medical Genetics, 2016. **53**(6): p. 419-425.
11. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. Hum Mol Genet, 2020. **29**(13): p. 2218-2239.
12. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. Genetics, 2017. **205**(1): p. 221-+.

13. De Keersmaecker, K., S.O. Sulima, and J.D. Dinman, *Ribosomopathies and the paradox of cellular hypo-to hyperproliferation*. *Blood*, The Journal of the American Society of Hematology, 2015. **125**(9): p. 1377-1382.
14. Kampen, K.R., et al., *Hallmarks of ribosomopathies*. *Nucleic Acids Research*, 2019. **48**(3): p. 1013-1028.
15. Schmid, M. and T.H. Jensen, *The Nuclear RNA Exosome and Its Cofactors*. *Adv Exp Med Biol*, 2019. **1203**: p. 113-132.
16. Ben-Shem, A., et al., *The structure of the eukaryotic ribosome at 3.0 Å resolution*. *Science*, 2011. **334**(6062): p. 1524-9.
17. Anttonen, A.-K., et al., *ZNHIT3 is defective in PEHO syndrome, a severe encephalopathy with cerebellar granule neuron loss*. *Brain*, 2017. **140**(5): p. 1267-1279.
18. Armistead, J., et al., *Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome*. *Am J Hum Genet*, 2009. **84**(6): p. 728-39.
19. Fraga de Andrade, I., C. Mehta, and E.H. Bresnick, *Post-transcriptional control of cellular differentiation by the RNA exosome complex*. *Nucleic Acids Res*, 2020. **48**(21): p. 11913-11928.
20. Dai, X. and M. Zhu, *Coupling of Ribosome Synthesis and Translational Capacity with Cell Growth*. *Trends Biochem Sci*, 2020. **45**(8): p. 681-692.
21. Pillon, M.C., Y.H. Lo, and R.E. Stanley, *IT'S 2 for the price of 1: Multifaceted ITS2 processing machines in RNA and DNA maintenance*. *DNA Repair (Amst)*, 2019. **81**: p. 102653.
22. Fang, F., S. Phillips, and J.S. Butler, *Rat1p and Rai1p function with the nuclear exosome in the processing and degradation of rRNA precursors*. *Rna*, 2005. **11**(10): p. 1571-8.
23. Selmer, M., et al., *Structure of the 70S ribosome complexed with mRNA and tRNA*. *Science*, 2006. **313**(5795): p. 1935-42.
24. Khatter, H., et al., *Structure of the human 80S ribosome*. *Nature*, 2015. **520**(7549): p. 640-5.
25. Fox, M.J., et al., *The Exosome Component Rrp6 Is Required for RNA Polymerase II Termination at Specific Targets of the Nrd1-Nab3 Pathway*. *PLOS Genetics*, 2015. **11**(2): p. e1004999.
26. Kilchert, C., S. Wittmann, and L. Vasiljeva, *The regulation and functions of the nuclear RNA exosome complex*. *Nature Reviews Molecular Cell Biology*, 2016. **17**(4): p. 227-239.

27. Zinder, J.C. and C.D. Lima, *Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors*. *Genes & Development*, 2017. **31**(2): p. 88-100.
28. Kiss, T., *Small Nucleolar RNAs: An Abundant Group of Noncoding RNAs with Diverse Cellular Functions*. *Cell*, 2002. **109**(2): p. 145-148.
29. Dupuis-Sandoval, F., M. Poirier, and M.S. Scott, *The emerging landscape of small nucleolar RNAs in cell biology*. *Wiley interdisciplinary reviews. RNA*, 2015. **6**(4): p. 381-397.
30. Huang, Z.-h., et al., *snoRNAs: functions and mechanisms in biological processes, and roles in tumor pathophysiology*. *Cell Death Discovery*, 2022. **8**(1): p. 259.
31. Sloan, K.E., et al., *Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function*. *RNA biology*, 2017. **14**(9): p. 1138-1152.
32. Kufel, J. and P. Grzechnik, *Small Nucleolar RNAs Tell a Different Tale*. *Trends in Genetics*, 2019. **35**(2): p. 104-117.
33. Han, C., et al., *Chromatin-associated orphan snoRNA regulates DNA damage-mediated differentiation via a non-canonical complex*. *Cell Reports*, 2022. **38**(13).
34. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases*. *Cell*, 1997. **91**(4): p. 457-466.
35. Mitchell, P., et al., *The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases*. *Cell*, 1997. **91**(4): p. 457-66.
36. Sterrett, M.C., et al., *A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex*. *Rna*, 2021. **27**(9): p. 1046-1067.
37. Fernández-Pevida, A., D. Kressler, and J. de la Cruz, *Processing of preribosomal RNA in *Saccharomyces cerevisiae**. *Wiley Interdiscip Rev RNA*, 2015. **6**(2): p. 191-209.
38. Henras, A.K., et al., *An overview of pre-ribosomal RNA processing in eukaryotes*. *Wiley Interdiscip Rev RNA*, 2015. **6**(2): p. 225-42.
39. Lafontaine, D.L., *Noncoding RNAs in eukaryotic ribosome biogenesis and function*. *Nat Struct Mol Biol*, 2015. **22**(1): p. 11-9.
40. Lestrade, L. and M.J. Weber, *snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs*. *Nucleic Acids Res*, 2006. **34**(Database issue): p. D158-62.
41. Karbstein, K., *Quality control mechanisms during ribosome maturation*. *Trends Cell Biol*, 2013. **23**(5): p. 242-50.

42. Lafontaine, D.L., *A 'garbage can' for ribosomes: how eukaryotes degrade their ribosomes*. Trends Biochem Sci, 2010. **35**(5): p. 267-77.
43. Rodríguez-Galán, O., et al., *Immature large ribosomal subunits containing the 7S pre-rRNA can engage in translation in Saccharomyces cerevisiae*. RNA Biol, 2015. **12**(8): p. 838-46.
44. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. J Neurol, 2013. **260**(7): p. 1866-70.
45. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. Nature Communications, 2014. **5**.
46. Burns, D., et al., *A recessive mutation in EXOSC9 causes abnormal RNA metabolism resulting in a novel form of cerebellar hypoplasia/atrophy with early motor neuronopathy*. Neuromuscul Disord, 2017. **27**: p. S38.
47. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
48. Eggens, V.R.C., et al., *EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations*. Orphanet Journal of Rare Diseases, 2014. **9**(1): p. 23.
49. Halevy, A., et al., *Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia*. J Neurol, 2014. **261**(11): p. 2165-9.
50. Rudnik-Schoneborn, S., et al., *Pontocerebellar hypoplasia type 1: clinical spectrum and relevance of EXOSC3 mutations*. Neurology, 2013. **80**(5): p. 438-46.
51. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
52. Somashekar, P.H., et al., *Bi-allelic missense variant, p.Ser35Leu in EXOSC1 is associated with pontocerebellar hypoplasia*. Clin Genet, 2021. **99**(4): p. 594-600.
53. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nat Genet, 2012. **44**(6): p. 704-8.
54. Yang, X., et al., *Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis*. Human Molecular Genetics, 2019. **29**(4): p. 541-553.
55. van Dijk, T., et al., *What's new in pontocerebellar hypoplasia? An update on genes and subtypes*. Orphanet journal of rare diseases, 2018. **13**(1): p. 92-92.

56. Fraga de Andrade, I., C. Mehta, and E.H. Bresnick, *Post-transcriptional control of cellular differentiation by the RNA exosome complex*. Nucleic acids research, 2020. **48**(21): p. 11913-11928.
57. Laffleur, B., et al., *RNA exosome drives early B cell development via noncoding RNA processing mechanisms*. Sci Immunol, 2022. **7**(72): p. eabn2738.
58. Lloret-Llinares, M., et al., *The RNA exosome contributes to gene expression regulation during stem cell differentiation*. Nucleic acids research, 2018. **46**(21): p. 11502-11513.

Appendix I. Generating a CRISPR/Cas9 toolkit to introduce RNA exosomopathy-linked missense mutations in *S. cerevisiae* genes.

The following encompasses the work of a SIRE student mentee, Jennifer Dean, who worked on this project for the school year 2021/2022. Text incorporates elements from Dean's Fall 2021 BIOL499 Progress Report.

A1.1 Abstract

The RNA exosome is an essential multi-subunit ribonuclease complex involved in the degradation and processing of RNAs. Missense mutations in genes that encode the subunits of the RNA exosome cause tissue specific diseases called RNA exosomopathies, like pontocerebellar hypoplasia and the novel syndrome SHRF. Assessing the consequences of these exosomopathy- linked amino acid substitutions in yeast has provided pivotal information about the molecular defects that could underlie patient pathology. However, it's unclear whether the mechanism of mutation insertion provides the most accurate *in-vivo* system. In previous studies, missense mutations were incorporated by plasmid shuffle assays, introduction exogenous expressing variants modeling the desired RNA exosome mutation into yeast strains with the endogenous wildtype gene knocked out. These plasmids are designed to be as genomic as possible, however since the mutations are not integrated in the genome, results may be less accurate given there could be plasmid copy differences between cells. To address this issue, we aimed to successfully introduce RNA exosomopathy mutations endogenously using the CRISPR/Cas9 system to alter the genomic sequence of the homologous yeast RNA exosome genes. Here we present our protocol from generating the CRISPR/Cas9 toolkit to screening potential clones. Our CRISPR/Cas9 toolkit is modeled after the gap-repair CRISPR/Cas9 protocol developed by the Ellis lab.

A1.2 Introduction

The processing and degradation of nearly every class of RNA in eukaryotes is carried out by the RNA exosome, an essential, multi-subunit ribonuclease complex. The RNA exosome is evolutionarily conserved, composed of a three-subunit cap (EXOSC1/2/3), a central six-subunit ring (EXOSC4/5/6/7/8/9), and a catalytically active ribonuclease subunit, DIS3 [1, 2] (Figure 1A). Recently, multiple tissue-specific human diseases, classified as “RNA exosomopathies”, have been linked to missense mutations in conserved regions of RNA exosome subunit genes [3-9]. Such diseases include neurodevelopmental disorders (EXOSC5 MIM *606492), pontocerebellar hypoplasia (EXOSC3, MIM #614678; EXOSC8, MIM #616081: and EXOSC9, MIM #618065) and a novel syndrome called SHRF (Short stature, Hearing loss, Retinitis pigmentosa and Faicies) with patients presenting with a myriad of phenotypes that include intellectual disability, neurodegeneration, retinitis pigmentosa, and others (EXOSC2, MIM #617763).

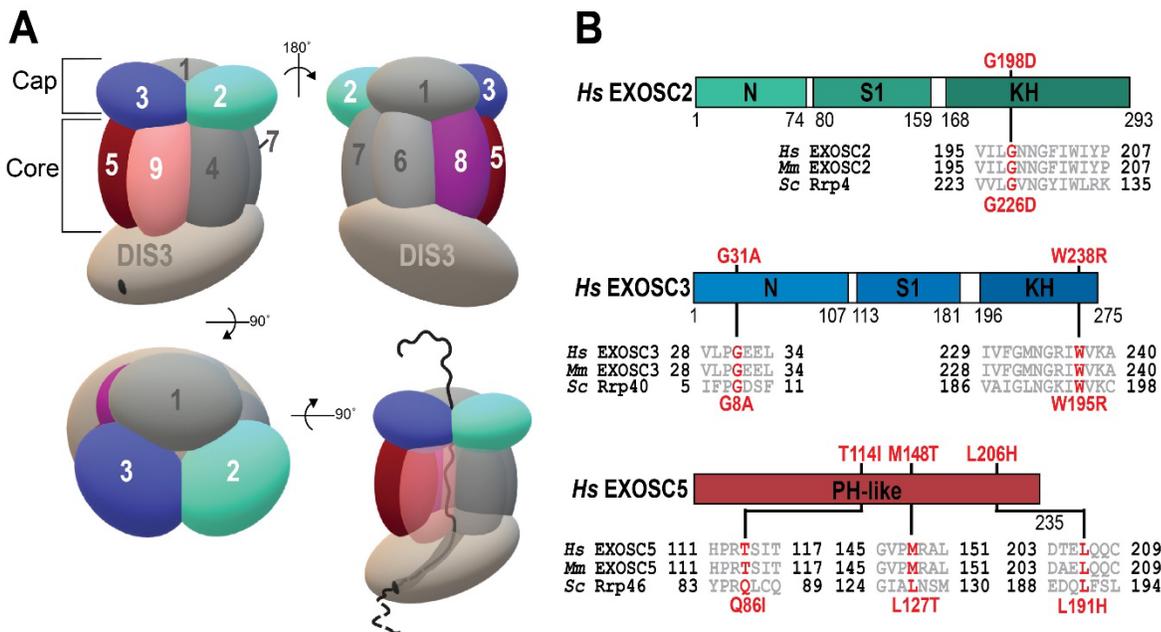


Figure 1. The RNA exosome is a multi-subunit complex that degrades/processes nearly every class of RNA. A) Diagram of the RNA exosome showing the three SI/KI cap subunits (EXOSC1/2/3), six PH-like domain core subunits (EXOSC4/5/6/7/8/9), and the catalytic DIS3 subunit. RNA is threaded through the cap and core to the catalytic DIS3 subunit. The subunits that are linked to RNA exosomopathy diseases are colored- EXOSC2 in teal (MIM #617763); EXOSC3 in blue (MIM #614678); EXOSC5 in red (MIM *606492); EXOSC8 in purple (MIM #616081); EXOSC9 in pink (MIM #618065) B) Domain map of the human EXOSC2, EXOSC3 and EXOSC5 subunits with sequence alignment across different eukaryotes. Highlighted in red are the RNA exosomopathy-linked amino acid substitutions found in patients (above) and their corresponding substitutions in the yeast orthologs (below).

These RNA exosomopathy mutations result in single amino acid changes in highly conserved regions of the subunits (Figure 1B). Given the disease variety and tissue-specificity, we predict that these amino acid changes result in distinct *in vivo* consequences that impact the function of the RNA exosome differently. Previous work has characterized the RNA exosomopathy mutations identified in *EXOSC2*, *EXOSC3* and *EXOSC5* by modeling the mutations in the *Saccharomyces cerevisiae* orthologs, *RRP4*, *RRP40* and *RRP46* [8, 10-12]. These studies have shown that these different mutations have functional and molecular consequences when modeled in *S. cerevisiae*. Specifically, the *EXOSC2 G198D*, *EXOSC3 W238R*, and *EXOSC5 L206H* mutations, corresponding to *rrp4-G226D*, *rrp40-W195R* and *rrp46 L191H* in the yeast system, result in defects in cell growth, cause an increase in steady state levels of RNA exosome targets, and impact the integrity of the RNA exosome complex. However, within all of these previous studies, the modeled RNA exosomopathy mutations have been expressed on exogenous plasmids in yeast cells that have the wildtype exosome subunit gene of interest knocked out through standard plasmid shuffle assays. While this method has proven useful in the study of RNA exosome mutations [8, 10, 11], we want to introduce these mutations in the endogenous yeast *RRP* genes to better assess the molecular and functional consequences resulting from these disease-linked amino acid changes.

To do this, we employed the CRISPR/Cas9 system to introduce our mutation of interest by editing the yeast genome directly. The CRISPR/Cas9 system is a bacterial immunological defense mechanism that has been adapted for genomic engineering in molecular biology [13]. The Cas9 enzyme can “cut” DNA at a specific target sequence located via a guide RNA, resulting in double strand breaks (DSB’s). It has been shown these DSB’s can be efficiently repaired via homologous recombination with donor DNA in *S. cerevisiae* [14], thus making the CRISPR/Cas9 an effective gene editing system to integrate mutations into endogenous yeast genes. We generated a CRISPR/Cas9 toolkit designed by the Ellis lab [15] to introduce RNA exosomopathy disease-linked mutations into corresponding yeast *RRP* genes. We then employed our toolkit through a workflow that resulted in successfully edited *RRP4* and *RRP40* genes that modeled RNA exosomopathy *EXOSC2* and *EXOSC3* missense mutations.

Within the CRISPR/Cas9 toolkit, we utilize a single guide RNA (sgRNA) construct that has a fusion of a 5' HDV ribozyme and a yeast tRNA structure which increases the stability of the sgRNA in cells [16]. This construct is included in the Ellis Lab Cas9 and sgRNA expression toolkit, which also includes a Cas9 ORF optimized for the *S. cerevisiae* system. Additionally, the toolkit employs a method of identifying yeast cells proficient in gap repair, sub-selecting cells that have increased competency of homologous recombination [17] and thus increasing our chances for efficient homologous recombination. We also designed asymmetric, single stranded donor oligonucleotides (ssODN) to introduce the RNA exosomopathy missense mutations as ssODN have increased efficiency in homology-directed repair in eukaryotic systems [18]. We utilized Benchling, an online platform, to design our guide RNAs and ssODNs. The following will describe our CRISPR/Cas9 toolkit, our genome editing workflow, and, lastly, present data of our successfully edited *rrp4* and *rrp40* mutant yeast cells employing our methodology.

A1.3 CRISPR/Cas9 Toolkit and Workflow

Our workflow is represented below in Figure 2:

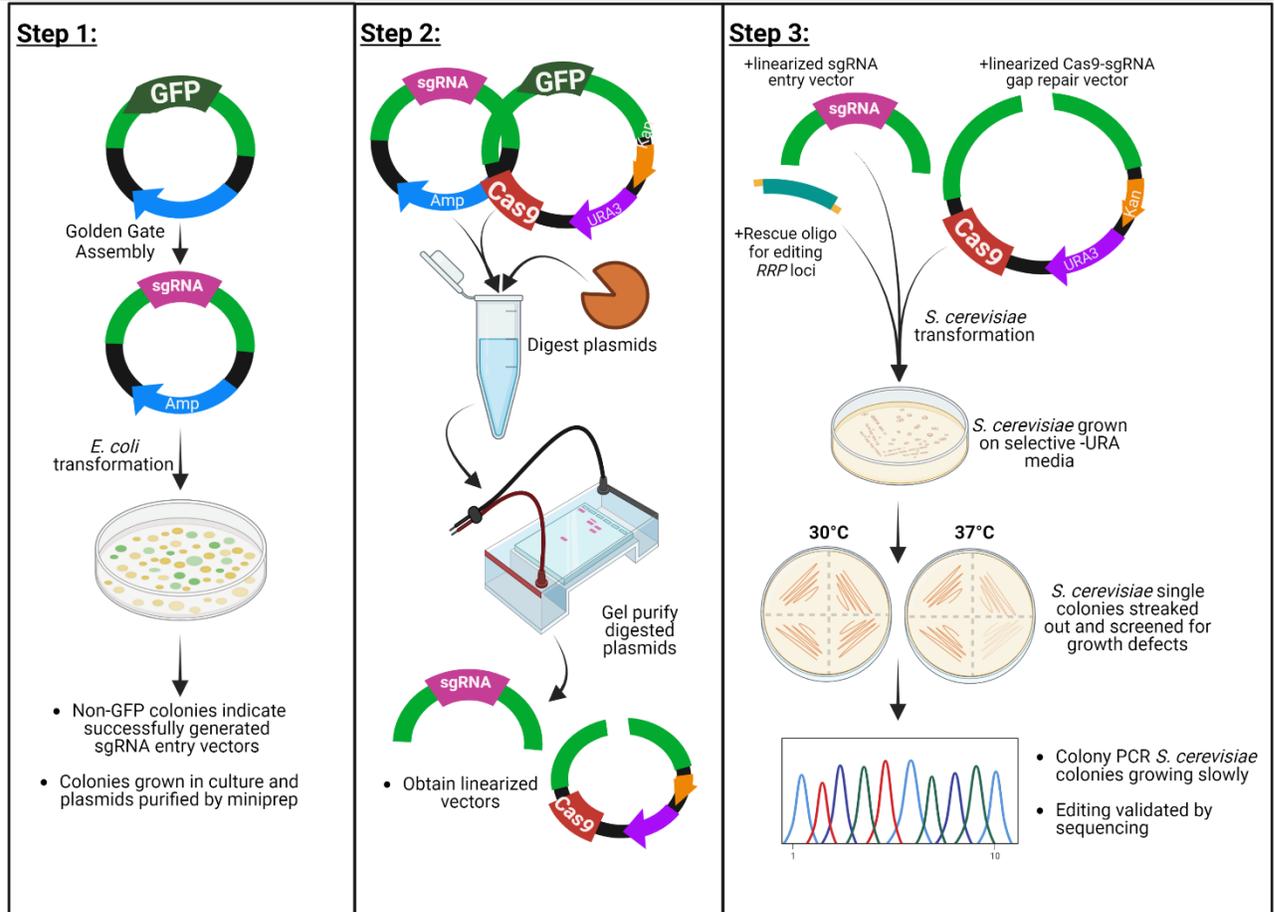


Figure 2. Workflow of CRISPR/Cas9 gene editing for *rrp* genes.

Steps 1 and 2 comprise of building the CRISPR/Cas9 toolkit, Step 3 employs the designed toolkit and screens for potentially edited yeast colonies.

STEP 1& 2: BUILDING OUR CRISPR/CAS9 TOOLKIT (Design by Ellis Lab

<https://benchling.com/pub/ellis-crispr-tools> [15])

The toolkit consists of:

- An sgRNA entry vector that contains the sgRNA sequence
- A Cas9-sgRNA expression vector with a selection marker
- Two ssODNs (or “rescue oligo”); one that contains the mutant variant sequence and one that contains just the wildtype variant.
 - These ssODNs also contained silent mutations within the PAM sites in order to make edited clones resistant to CRISPR/Cas9 targeting and cutting.

Table 1 summarizes the vectors and oligonucleotides we used and generated for these experiments.

Design of sgRNAs and generating the sgRNA entry vector

Our sgRNAs were designed using the online informatics platform Benchling (<https://help.benchling.com/hc/en-us>). Yeast DNA sequences of interest were uploaded to the platform, and the desired mutation sites were located. We limited the range for targeting by the sgRNA to 20 bases upstream or downstream the mutation locus. From this sequence, the Benchling platform identified a different gRNA option, and we then selected the “best” option considering the on-target and off-target scores as calculated by Benchling (higher scores were considered “better”). We also aimed to have sgRNAs with corresponding PAM sequences in frame or the first 2 nucleotides of the PAM site had to be the second and third nucleotides of a codon (i.e. NCG|GNN) of the gene in order to more easily generate ssODNs with silent mutations in the PAM sites (see “*Design of ssODNs*” for more details).

Thermocycler protocol

1. 42°C for 2min
2. 16°C for 5min
3. Repeat steps 1-2 (10x)
4. 60°C for 10min
5. 80°C for 10min

Half the reaction mixture was transformed into *E.coli* and we selected for non-GFP colonies.

Figure 3 is representative of the efficiency observed in the Golden Gate Assembly protocol and the BsmBI digest. Non-GFP colonies were inoculated in selective media and miniprepmed. *Table 1* summarizes the vectors generated containing *RRP* targeting sgRNAs using this method.

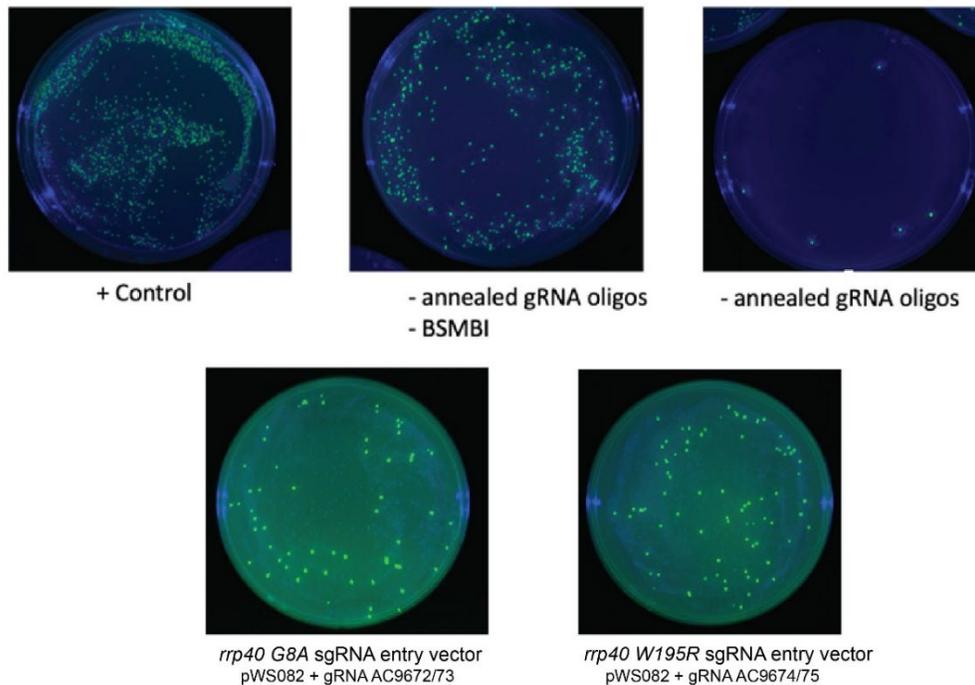


Figure 3. *E. coli* transformation of reaction mixtures from the small fragment Golden Gate assembly shows relatively efficient generation of sgRNA vectors with annealed sgRNA oligonucleotides. Positive control with just the intact sgRNA entry vector plasmid shows high GFP expression (labeled “+ Control”). A Golden Gate assembly reaction performed with the empty sgRNA entry vector but without BsmBI or annealed oligos shows relatively high GFP expression (labeled “-annealed gRNA oligos, -BSMBI”). Similarly, Golden Gate assembly reaction performed with the empty sgRNA entry vector without annealed oligos on the right shows the lowest GFP expression (labeled “-annealed gRNA oligos”). Two plates presented on the second row are representative of two successful small fragment Golden Gate assembly reactions performed with annealed sgRNAs to target different *rrp40* loci. The *rp4 G8A* and *rrp40-W195R* plates reveal some GFP expression but many colonies a non-GFP, indicating that sgRNA was successfully integrated in the sgRNA entry vector.

Digest and linearizing sgRNA and Cas9 vectors for gap repair selection

With validated sgRNA containing vectors, plasmids are linearized by EcoRV to create a linear sgRNA expression cassette (Figure 4A). Simultaneously, the Cas9-sgRNA expression vector is linearized with a BsmBI restriction digest, removing a GFP cassette (Figure 4A). Flanking the BsmBI cut sites in the Cas9-sgRNA vector are 500bp homology arms to the sequence surrounding the linearized sgRNA vector. These homology arms will direct gap repair by homologous recombination within *S. cerevisiae* when these pieces are transformed into the cells (Figure 4B).

The selective pressure for gap repair by the *S. cerevisiae* cells will be plating the cells on -URA minimal media. Digested products were run on a 1% agarose gel and gel purified using the Qiagen gel extraction protocol.

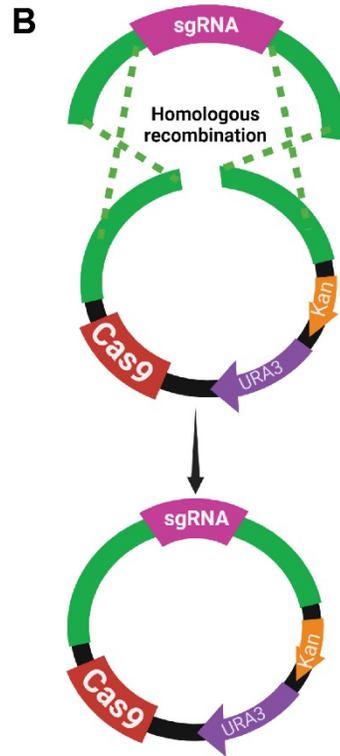
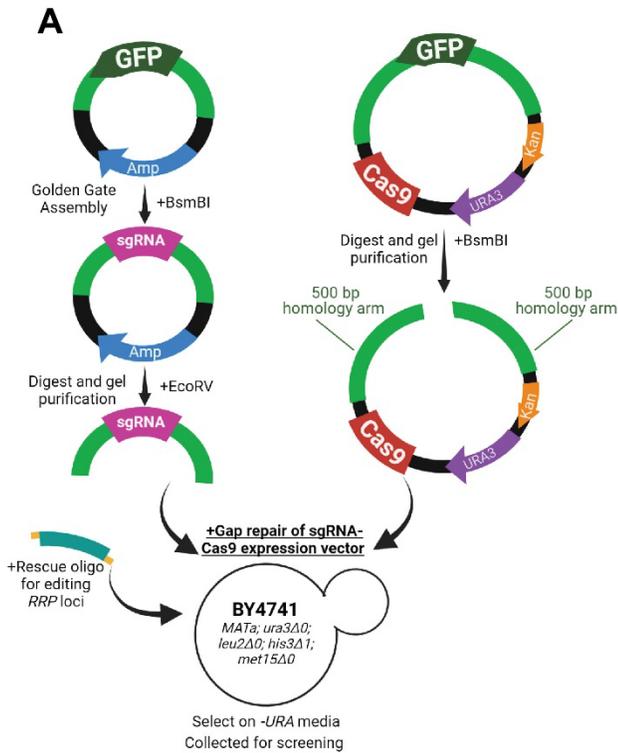


Figure 4. A schematic of building our CRISPR/Cas9 toolkit. (A) The sgRNA entry vector is generated using Golden Gate assembly. The sgRNA vector and the Cas9-sgRNA vectors are digested and purified. These linearized pieces are transformed into *BY4741* *S. cerevisiae* where they will be homologously recombined through gap repair. Cells are selected on -URA minimal media, to select for the recombined Cas9-sgRNA plasmid. The rescue oligo/ssODN is also transformed in the yeast in order to edit the *rrp* locus. (B) Cartoon depiction of gap repair of the Cas9-sgRNA plasmid. The *S. cerevisiae* will homologously recombine the linearized sgRNA vector and the linearized Cas9-sgRNA plasmid through the overlapping 500bp homology arms present on both

Design of ssODNs

We based our rescue oligonucleotide design on the findings of Richardson *et al.* [18] (Figure 5).

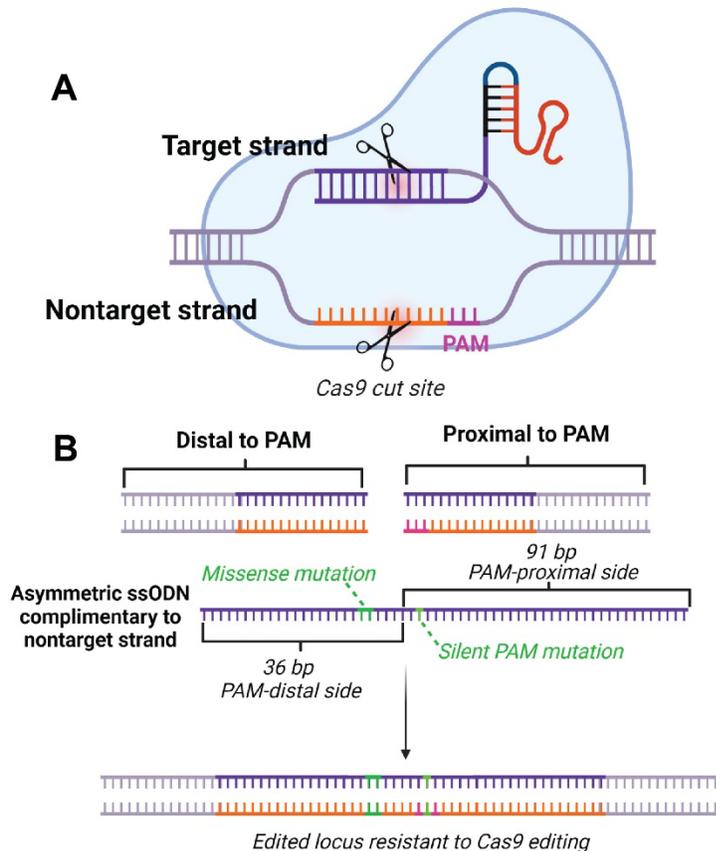


Figure 5. Schematic of our ssODN design. (A) The sgRNA-Cas9 ribonuclease complex is directed to a target locus. The sgRNA binds to the target strand and the PAM site on the complementary nontarget strand is recognized by Cas9. The enzyme makes a double stranded break three nucleotides upstream of the PAM site. (B) The asymmetric ssODN flanks the double stranded break, 36 base pairs on the PAM-distal side and 91 base pairs on the PAM-proximal side. The ssODN contains the desired mutation and a silent mutation in the PAM site to make edited clones resistant to iterative editing by Cas9.

The Cas9-sgRNA complex is guided to a target based on the complementarity of the sgRNA sequence. The DNA strand that the sgRNA complements to is considered the “Target strand” (Figure 5A). The other DNA strand is the “Non target” strand (Figure 5A). The Cas9 enzyme then “cuts” three nucleotides upstream of the PAM site, first on the “Nontarget strand” and then secondly on the “Target strand”. As a result of this double strand break, there is a PAM-proximal and a PAM-distal side of the DNA breakpoint (Figure 5B).

Richardson *et al.* found that generating an asymmetric single stranded donor DNA enhanced homologous recombination of a Cas9 induced double stranded break

[18]. The size parameters reported as most efficient was a ssODN that spans 36 bases of the PAM-distal side and 91 bases of the PAM-proximal side [18]. Additionally, having the ssODN

complementary to the “Nontarget strand” of DNA was more effective than a ssODN complementary to the “Target” strand [18].

Lastly, we also wanted our ssODN to harbor silent mutations in the PAM sites, in addition to the exosomopathy mutations, in order to make edited cells resistant to further targeting by the CRISPR/Cas9 system. As our CRISPR/Cas9 toolkit utilizes a plasmid that expresses the enzyme, we knew that during our workflow cells would be expressing Cas9 and sgRNA for multiple days. If the cells were edited, we wanted to ensure that there wouldn't be secondary rounds of targeting and DNA “cutting” which could introduce errors and secondary mutations. One way to alleviate this multi-round Cas9 targeting is to mutate the PAM site. As described in “*Design of sgRNAs and generating the sgRNA entry vector*”, we tried to select gRNAs that were in frame or the first 2 nucleotides of the PAM site had to be the second and third nucleotides of a codon (i.e. NCG|GNN). This would allow for us to introduce silent mutations in our ssODN that would remove the PAM site. For example:

- PAM site in frame: CGG codon encodes for Arginine ; Alternatively could use CGA/C/U to encode for Arginine
- PAM site NCG|GNN: The second “G” of the “CGG” PAM site is in the wobble position and therefore can be more easily changed to generate a “C A/U/C G” sequence

It's important to note that in both situations, we used codon usage tables to try and find alternate codons that were comparable to one another when introducing these silent PAM site mutations!

With every ssODN, we had two varieties-one that introduced our desired RNA exosomopathy mutation and one that only had the PAM site mutation. The latter was used to generate a “wildtype” control line.

In summary, we used the following constraints in designing our ssODN:

1. Complementary sequence of the “Nontarget” strand
2. Asymmetric design, spanning 36 bases PAM-distal and 91bases PAM-proximal
3. Silent mutation at the PAM site

STEP 3: UTILIZING CRISPR/CAS9 TOOLKIT AND SCREENING YEAST

Transforming yeast with CRISPR/Cas9 toolkit components

We used a High-efficiency Transformation of Yeast protocol to transform our desired DNA mixtures into

yeast cells. For genome editing we had these three components at the following concentrations:

- 100 ng of linearized Cas9-sgRNA gap repair vector
- 200 ng of digested sgRNA vector
- > 2 µg of ssODNs (either the “Wildtype” or “Mutant” ssODN)
 - A “Wildtype” ssODN containing only the PAM silent mutation and
 - A “Mutant” ssODN, containing the PAM mutation and the modeled RNA exosomopathy missense mutation)

Transformations were plated on -URA minimal media to select for gap repaired plasmids and, thus, cells competent for homologous recombination. Figure 6 shows the efficiency we observed in transformation and gap repair.

After two days of growth, multiple colonies from both the mutant oligonucleotide and wildtype oligonucleotide transformation reactions were struck onto YEPD and grown at 30°C and 37°C to initially screen for any growth defects. We predicted that edited cells that endogenous express the RNA exosomopathy mutations would have slower growth at 37°C given previous studies showing that cells exogenously expressing mutated *rrp* variants had functional consequences [8, 10, 11]. The colonies that showed slower growth at 37°C after ~2 days were further screened through DNA sequencing. Additionally, we sequenced colonies from the wildtype oligonucleotides transformation that showed no growth defect at 37°C to confirm that they only contained the silent PAM mutation. Colonies that showed editing at the *rrp* locus were collected

and used for further studies. *Table 2* summarizes the frozen stocks of *rrp4* and *rrp40* mutants generated using this method and utilized in the following Results section.

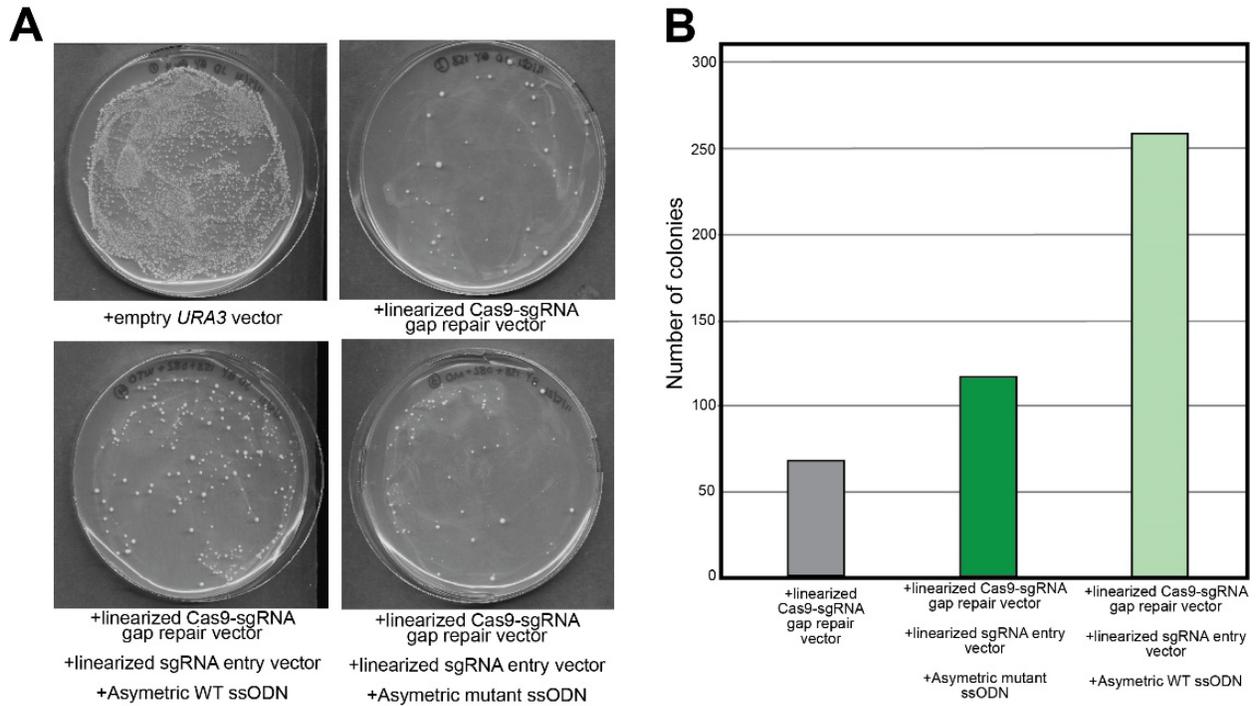


Figure 6. Representative yeast transformation plates show efficient gap repair. A) Images of transformant plates. Top left; transformation with an intact *URA3* plasmid as a positive control. Top right; transformation with the linearized Cas9-sgRNA vector shows fewer colonies, suggesting successful digest of the plasmid. Bottom left; transformation with the linearized sgRNA and Cas9-sgRNA vectors and a wildtype rescue oligonucleotide (WT ssODN). Bottom right; transformation with the linearized sgRNA and Cas9-sgRNA vectors and a mutant rescue oligonucleotide (mutant ssODN). B) Graphical representation of the number of colonies on the transformant plates.

A1.4 Results: Generating and assessing *rrp4-G226D* and *rrp40-W195R* mutant cells

We first focused on generating *rrp4-G226D* and *rrp40-W195R* mutant cells. From previous work done characterizing the *rrp4-G226D* and *rrp40-W195R* variants [10, 11], we anticipated edited cells to endogenously express *rrp4-G226D* and *rrp40-W195R* would have similar growth phenotypes as those described. Therefore, we screened our transformants by selecting colonies and assessing growth at 30°C and 37°C (Figure 7 and Figure 8). From this initial screen we identified some potential *rrp4-G226D* edited colonies that showed growth defects at 37°C (Figure 7). We selected those potential *rrp4-G226D* colonies (labeled “Mutant”) as well as the corresponding CRISPR/Cas9 wildtype colonies (labeled “WT”) and performed a colony PCR to amplify the *RRP4* locus. Successfully amplified PCR products were sequenced and analyzed to determine which mutant colonies had truly both the PAM site mutation and the desired *rrp4-G226D* mutation. Mutant colonies M and S had edited *rrp4* loci that incorporated both the *rrp4-G226D* mutation and the silent PAM site mutation. Wildtype colonies C, F, M, P, and S had the edited

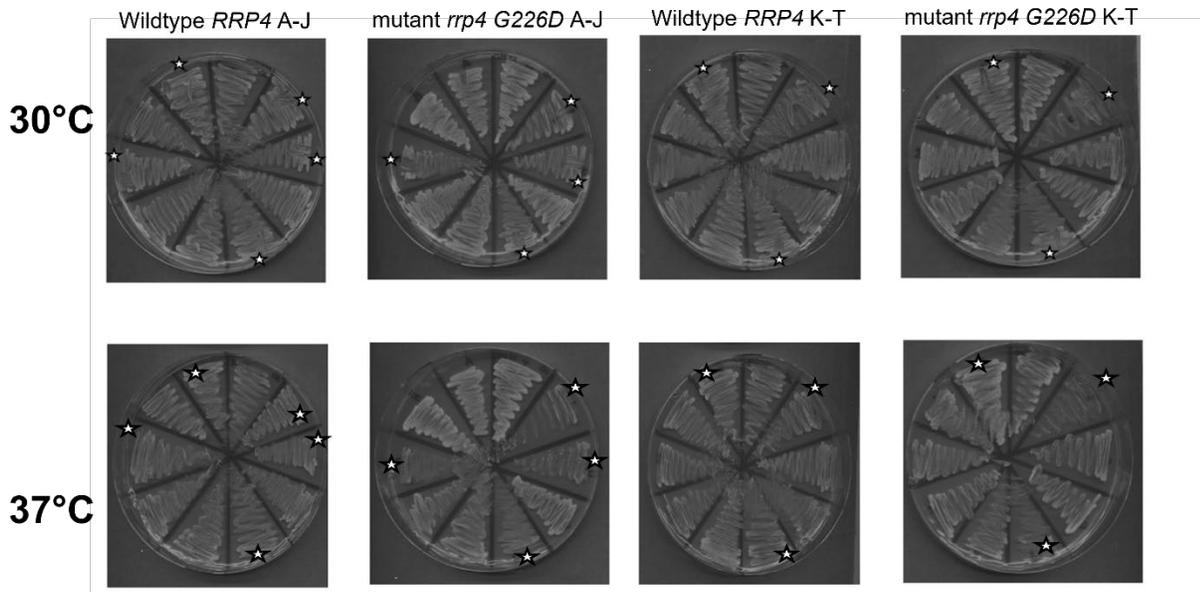


Figure 7. Initial screen of potential *rrp4-G226D* edited colonies. 20 colonies from both the mutant oligonucleotide and wildtype oligonucleotide transformation reactions were streaked out and grown at 30°C and 37°C. Potential *rrp4-G226D* colonies were identified by visible growth defects at 37°C. These potential colonies and corresponding edited wildtype colonies were selected for colony PCR and sequencing. Stars indicate selected *rrp4-G226D* edited colonies and *RRP4* edited colonies that were further assessed by sequencing.

PAM site change. Similarly, we identified some potential *rrp40-W195R* edited colonies that showed growth

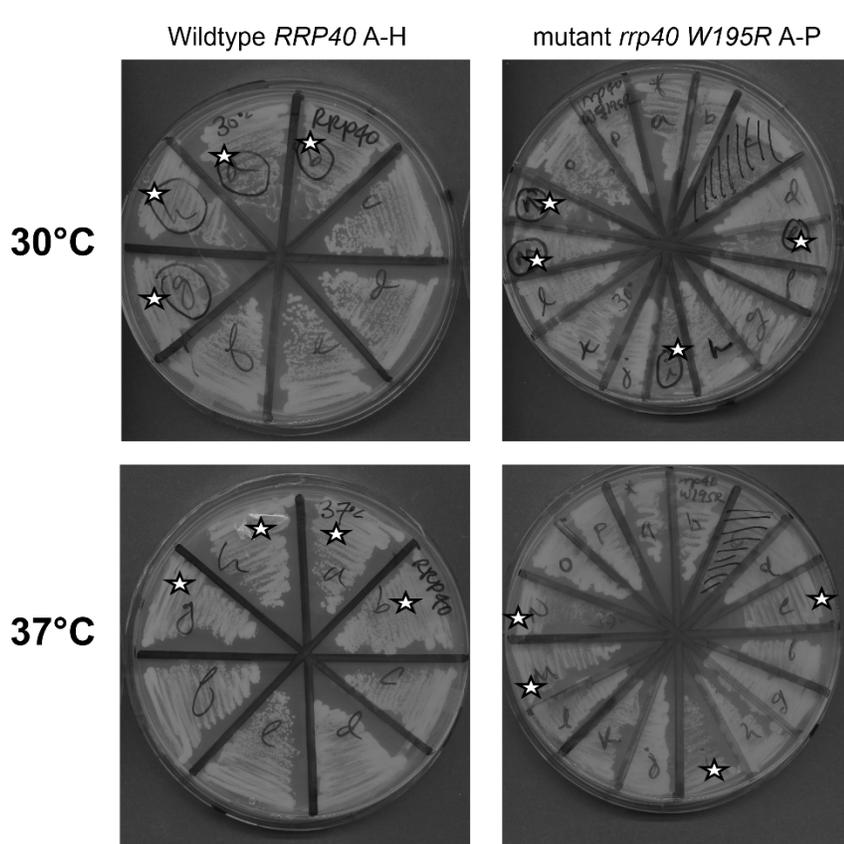


Figure 8. Initial screen of potential *rrp40-W195R* edited colonies. 16 colonies from the mutant oligonucleotide transformation and 8 colonies from the wildtype oligonucleotide transformation were streaked out and grown at 30°C and 37°C. Potential *rrp40-W195R* colonies were identified by visible growth defects at 37°C. These potential colonies and selected wildtype colonies were selected for colony PCR and sequencing. Stars indicate selected *rrp40-W195R* edited colonies and *RRP40* edited colonies that were further assessed by sequencing.

defects at 37°C (Figure 8). We selected those potential *rrp40-W195R* colonies, performed a colony PCR to amplify and sequenced the *RRP40* locus. Mutant colonies E, I, M and N had edited *rrp40* loci that incorporated both the *rrp40-W195R* mutation and the silent PAM site mutation. Wildtype colonies A, B, G, and H had the silent PAM site change.

To characterize these CRISPR edited *rrp* mutant cells, we performed a solid media growth assay on the edited wildtype and mutant M and S colonies (Figure 9). Results of

these spotting assays reveal growth defects for the edited *rrp4-G226D* mutants at 30°C and 37°C on YEPD plates (Fig 9A). These results are consistent with previous growth defect results from experiments utilizing exogenous plasmid mutation expression [10]. We also observed growth defects for the edited *rrp40-W195R* mutants at 37°C on YEPD, consistent with previous studies (Figure 9B) [11, 12]. Importantly, we do not see any growth defects in our wild type control *RRP4* and *RRP40* cells, suggesting that the PAM site mutations have no impact on cell growth and that we did not inadvertently introduce any deleterious off-target effects from the CRISPR/Cas9 machinery (Figure 9). We also compared the growth of our edited *rrp4-G226D* and *rrp40-W1955R* clones to the plasmid containing mutants that have previously been

published (Figure 10) [10, 11]. We performed liquid growth assays and calculated the doubling time of the mutant and wildtype cells. The edited *rrp4-G226D* clones have slower growth at both 30°C and 37°C, as suggested by the solid media assays, but we now can conclude that this defect is worse than that of the plasmid expressing *rrp4-G226D* mutant cells (*rrp4Δ; rrp4-G226D*) (Figure 10A). Similarly, we see that the *rrp40-W195R* edited clones have slower growth at 37°C compared to the plasmid expressing *rrp40-W195R* mutant cells (*rrp40Δ; rrp40-W195R*) (Figure 10B). Interestingly we also see a slight growth defect at 30°C with the liquid media assay for the *rrp40-W195R* cells that has not been observed before for this variant (Figure 10B).

To determine if the observed growth defects in the edited *rrp4 G226D* and *rrp40 W195R* cells are solely due to the *rrp* mutations, we transformed the *rrp4-G226D* and *rrp40-W195R* mutant clones with plasmids expressing wildtype *RRP4* and *RRP40*. We see rescue of the growth defects observed in the *rrp4-G226D* mutant clones with expression of the *RRP4* plasmid at both 30°C and 37°C (Figure 11A). Similarly, we see rescue of the growth defects observed in the *rrp40-W195R* mutant clones with expression of the *RRP40* plasmid at both 37°C (Figure 11B). These data further suggest that our CRISPR/Cas9 toolkit did not introduce any off-target effects and that the observed growth defects are due to the edits made at the *rrp* locus.

Lastly, we assessed the growth of our edited *rrp* mutant clones on different drug plates that disrupt different cellular pathways (Fig 12). Carreine impairs the cellular stress response/TOR signaling [19], formamide alters RNA metabolism [20], and hydroxyurea impairs DNA synthesis [21]. For comparison, we included *rrp4Δ; rrp4-G226D* and *rrp40Δ; rrp40-W195R* cells with the corresponding wild type controls in these solid media growth assays. As shown in Figure 12A, the *rrp4 G226D* clones show a similar sensitivity to hydroxyurea as the *rrp4Δ; rrp4-G226D* cells however show an increased sensitivity to both formamide and caffeine. In contrast, the edited *rrp40-W195R* cells do not show any sensitivity to these drugs (Figure 12B). This is consistent with the growth observed for the *rrp40Δ; rrp40-W195R*.

A1.5 Discussion

Our CRISPR/Cas9 toolkit and workflow successfully generated cells with endogenous *rrp4-G226D* and *rrp40-W195R* mutations. The methods used can be applied to generating other *rrp* mutations that can model RNA exosomopathy mutations, particularly those that have yet to be characterized. Furthermore, the design of this toolkit by the Ellis lab allows for iterative mutations within a singular experiment, suggesting we could generate combinatorial *rrp* mutants or introduce mutations in RNA exosome cofactor genes within the same cell line. In validating our methodology and characterizing our edited *rrp* mutant clones, we see stronger growth defects in the edited *rrp4-G226D* and *rrp40-W195R* cells compared to those previously reported [10, 11], suggesting that these mutations have more consequences when expressed endogenously. From our work presented here we show how we developed an efficient, accurate and fast system to introduce these RNA exosomopathy mutations endogenously in *S. cerevisiae*, giving us a broader toolset to assess the molecular and functional consequences resulting from these human disease mutations.

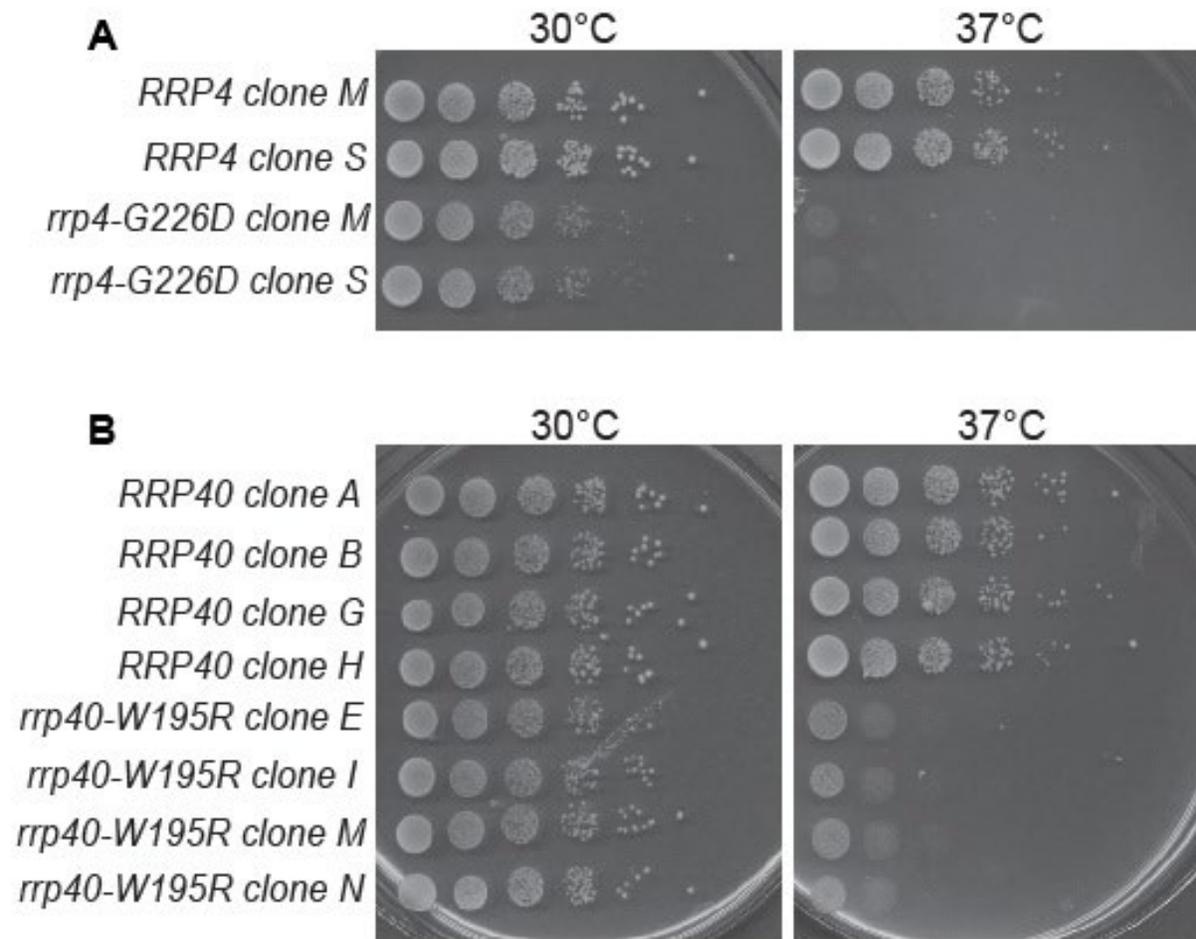


Figure 9. Solid media growth assay of identified CRISPR edited *rrp4-G226D* (A) and *rrp40-W195R* cells (B). Overnight cultures of edited clones were serially diluted and grown on YEPD media at 30°C and 37°C for two days.

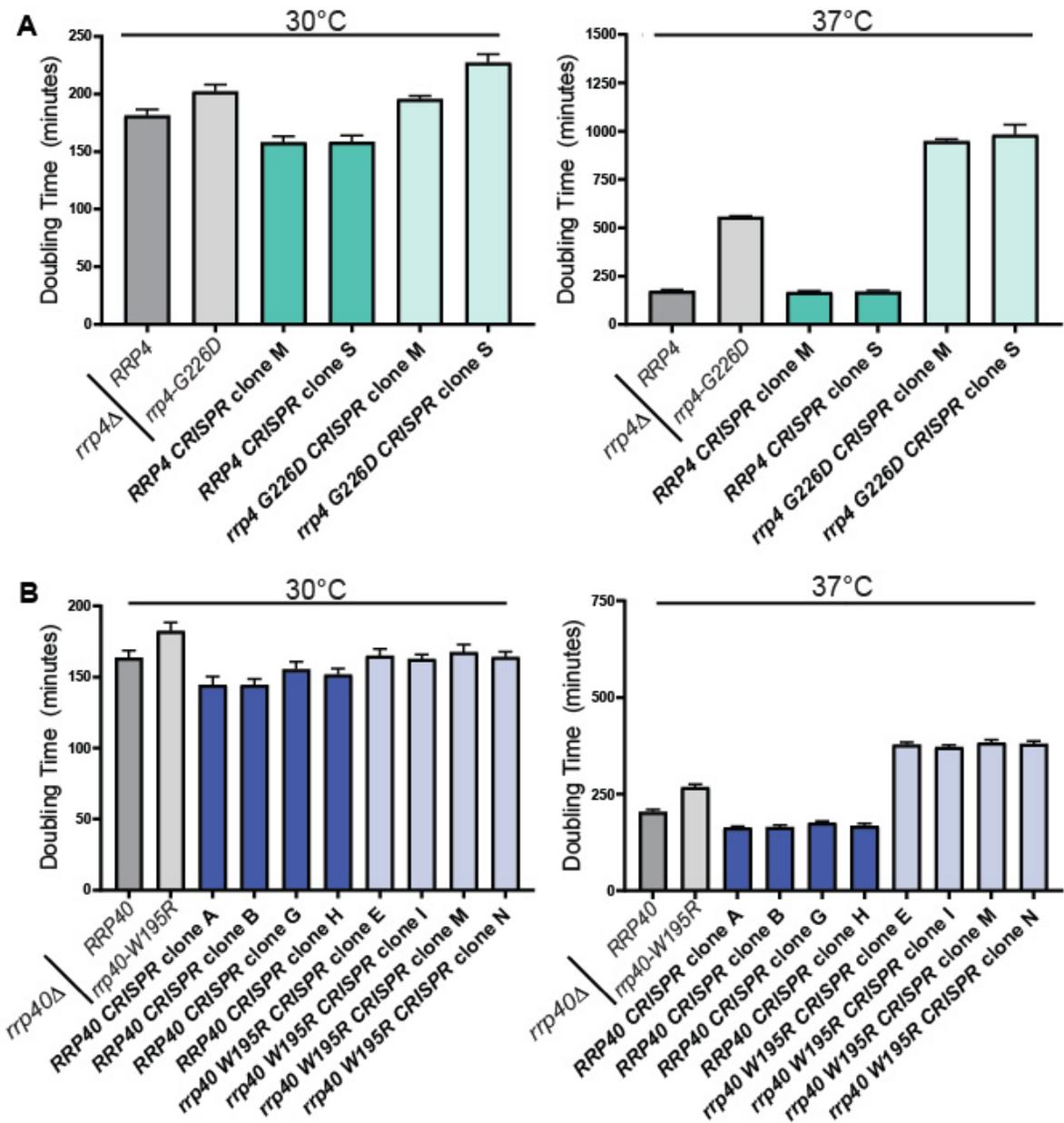


Figure 10. Liquid media growth assay of identified CRISPR edited *rrp4-G226D* (A) and *rrp40-W195R* cells (B). Overnight cultures of edited clones were diluted and grown in YEPD media at 30°C and 37°C. Doubling times were calculated and graphed. Included in the liquid growth assays were the *rrp4Δ*; *rrp4-G226D* and *rrp40Δ*; *rrp40-W195R* cells to compare defects of the plasmid expressing and edited *rrp* mutants.

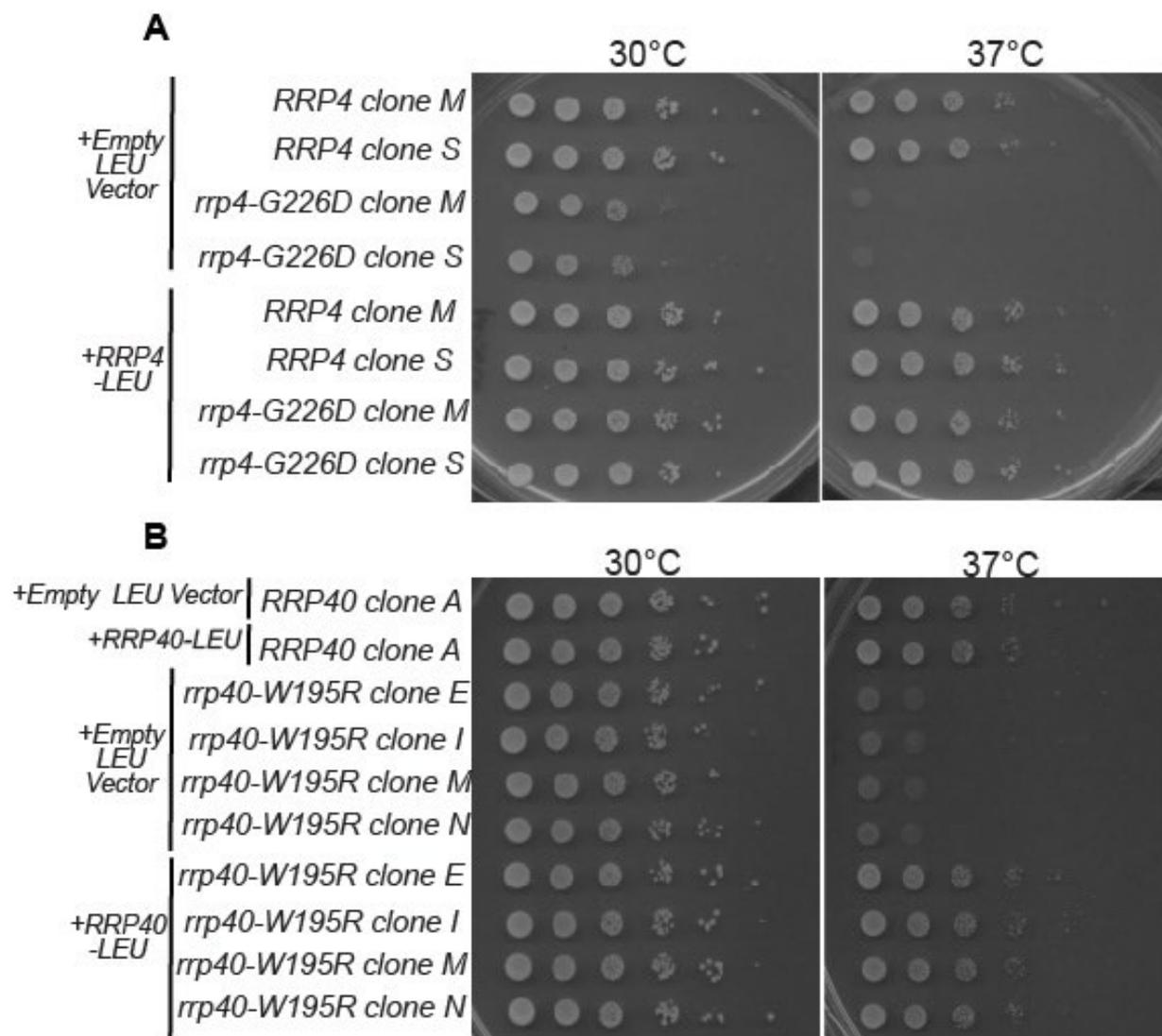


Figure 11. Rescue of CRISPR edited *rrp4-G226D* (A) and *rrp40-W195R* cells (B). Edited clones were transformed with plasmids expressing wild type *RRP4* or *RRP40*. Cells were grown to saturation, serially diluted and grown on selective media at 30°C and 37°C for two days.

Table 1. CRISPR/Cas9 Toolkit parts list

Description	Name	Sequence
sgRNA entry vector	<u>pWS082</u>	
Cas9-sgRNA gap repair vector	<u>pWS158</u>	
Rrp46 L191H CRISPR Ellis gRNA rev	AC9681	AAACGCTCCAGTAAACTGAACAATAA
Rrp46 L191H CRISPR Ellis gRNA fwd	AC9680	GACTTTATTGTTTCAGTTTACTGGAGC
Rrp46 L127T CRISPR Ellis gRNA rev	AC9679	AAACTAAATAGTATGTGCGCAAGCAA
Rrp46 L127TCRISPR Ellis gRNA fwd	AC9678	GACTTTGCTTGCGCACATACTATTTA
Rrp46 Q86I CRISPR Ellis gRNA rev	AC9677	AAACAAGACAGCTCTGCCAGATAAAA
Rrp46 Q86I CRISPR Ellis gRNA fwd	AC9676	GACTTTTTTATCTGGCAGAGCTGTCTT
Rrp40 W195R CRISPR Ellis gRNA rev	AC9675	AAACCGCACTTAACCCAGATCTTCAA
Rrp40 W195R CRISPR Ellis gRNA fwd	AC9674	GACTTTGAAGATCTGGGTAAAGTGCG
Rrp40 G8A CRISPR Ellis gRNA rev	AC9673	AAACTGGTGATAGCTTTCCTGTAGAA
Rrp40 G8A CRISPR Ellis gRNA fwd	AC9672	GACTTTCTACAGGAAAGCTATCACCA
ssODN Rrp4 G226D asymmetric HDR	AC9624	TTAAGAAACGGGATGTTTTGCCAAGTCCCGAGTTCATTA ATAGTGAGAGCCAAGAACCATACTCATAATTTGCCCGGC AACATAACAGTAGTTCTCGATGTCAATGGTTACATATGG TTAAGGAAAA
ssODN Rrp4 G226 WT asymmetric HDR	AC9623	TTAAGAAACGGGATGTTTTGCCAAGTCCCGAGTTCATTA ATAGTGAGAGCCAAGAACCATACTCATAATTTGCCCGGC

		AACATAACAGTAGTTCTCGGAGTCAATGGTTACATATGG TTAAGGAAAA
Rrp40 W195R mutant ssODN	AC9701	TTTGCTATATCCTTGAACGCTGCCGTGTCGTTTTTTTGAC AACACTCCATTATGGTTCTATAACAAGCTAAAGTGTTAG ATAATTCTTCGCACTTAACACGGATCTTCCCATTGAGACC AATGGCGA
Rrp40 W195R WT ssODN	AC9700	TTTGCTATATCCTTGAACGCTGCCGTGTCGTTTTTTTGAC AACACTCCATTATGGTTCTATAACAAGCTAAAGTGTTAG ATAATTCTTCGCACTTAACCCAGATCTTCCCATTGAGACC AATGGCGA
Rrp46 L191H mutant ssODN	AC9699	AATGGCTATACTAGACGAGGTGATATATTGTCCTGT ATTATTCTTCTTATATTTGTGACAAGTTCTTGACACTTTT GCTCACCTAGCTCCAGTAAACTGAAATGTTGATCTTCATT AAAATCAC
Rrp46 L191H WT ssODN	AC9698	AATGGCTATACTAGACGAGGTGATATATTGTCCTGT ATTATTCTTCTTATATTTGTGACAAGTTCTTGACACTTTT GCTCACCTAGCTCCAGTAAACTGAACAATTGATCTTCAT TAAAATCAC

Table 2. Designed sgRNA vectors with rrp sgRNAs

Description	Name	Sequence
Rrp46 L191H CRISPR Ellis gRNA fwd	MS162	pWS082 + gRNA AC9680/9681
Rrp46 L127T CRISPR Ellis gRNA rev	MS163	pWS082 + gRNA AC9678/9679
Rrp46 Q86I CRISPR Ellis gRNA rev	MS164	pWS082 + gRNA AC9677/9676
Rrp40 W195R CRISPR Ellis gRNA rev	MS165	pWS082 + gRNA AC9674/75

Rrp40 G8A CRISPR Ellis gRNA rev	MS166	pWS082 + gRNA AC9672/73
------------------------------------	-------	-------------------------

A1.6 References

1. LaCava, J. and S. Vanacova, *Correction to: The Eukaryotic RNA Exosome*. Methods Mol Biol, 2020. **2062**: p. C1-C4.
2. Morton, D.J., et al., *The RNA exosome and RNA exosome-linked disease*. Rna, 2018. **24**(2): p. 127-142.
3. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. J Neurol, 2013. **260**(7): p. 1866-70.
4. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. Nature Communications, 2014. **5**.
5. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. Am J Hum Genet, 2018. **102**(5): p. 858-873.
6. Di Donato, N., et al., *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*. Journal of Medical Genetics, 2016. **53**(6): p. 419-425.
7. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
8. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. Hum Mol Genet, 2020. **29**(13): p. 2218-2239.
9. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. Nature Genetics, 2012. **44**(6): p. 704-U134.
10. Sterrett, M.C., et al., *A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex*. Rna, 2021. **27**(9): p. 1046-1067.
11. Fasken, M.B., et al., *Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast*. Genetics, 2017. **205**(1): p. 221-+.
12. Gillespie, A., et al., *Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in S. cerevisiae*. RNA, 2017. **23**(4): p. 466-472.
13. Jiang, F. and J.A. Doudna, *CRISPR–Cas9 Structures and Mechanisms*. Annual Review of Biophysics, 2017. **46**(1): p. 505-529.
14. DiCarlo, J.E., et al., *Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems*. Nucleic Acids Res, 2013. **41**(7): p. 4336-43.

15. Shaw, W., Ellis Lab. *Quick and easy CRISPR engineering in Saccharomyces cerevisiae*. Available from: <https://benchling.com/pub/ellis-crispr-tools>.
16. Ryan, O.W., et al., *Selection of chromosomal DNA libraries using a multiplex CRISPR system*. eLife, 2014. **3**: p. e03703.
17. Horwitz, A.A., et al., *Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas*. Cell Syst, 2015. **1**(1): p. 88-96.
18. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA*. Nature Biotechnology, 2016. **34**(3): p. 339-344.
19. Kuranda, K., et al., *Investigating the caffeine effects in the yeast Saccharomyces cerevisiae brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways*. Mol Microbiol, 2006. **61**(5): p. 1147-66.
20. Hoyos-Manchado, R., et al., *RNA metabolism is the primary target of formamide in vivo*. Sci Rep, 2017. **7**(1): p. 15895.
21. Slater, M.L., *Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle*. J Bacteriol, 1973. **113**(1): p. 263-70.

Appendix II. Experimental Design to assess the impacts of pathogenic amino acid substitutions on RNA exosome integrity

The following encompasses the work of a SIRE student mentee, Will Ball, who worked on this project for the school year 2021/2022. Text incorporates elements from Ball's Spring 2022 BIOL499 research project and poster.

A1.1 Introduction

The processing and degradation of nearly every class of RNA in eukaryotes is carried out by the RNA exosome, an essential, multi-subunit ribonuclease complex. The RNA exosome is evolutionarily conserved, composed of a three-subunit cap (EXOSC1/2/3), a central six-subunit ring (EXOSC4/5/6/7/8/9), and a catalytically active ribonuclease subunit, DIS3 [1, 2] (Figure 1A). The genes that encode the RNA exosome were first identified in a screen for ribosomal RNA processing (*rrp*) mutants in *S. cerevisiae* [4]. Through genetics and biochemical assays, the authors of this seminal study identified that mutations in those *rrp* genes resulted in defective 3' maturation of 5.8S rRNA and that the encoded Rrp proteins were in complex with one another [4]. Since, several studies have resolved the structures of both the human and yeast RNA exosome, confirming a conserved organization of multiple non-catalytically active structural subunits (Figure 1B) [3, 5-9].

While much is known about the overall structure and the transcripts the RNA exosome targets, there is still much to be discovered about the integrity of the complex, such as association or dissociation

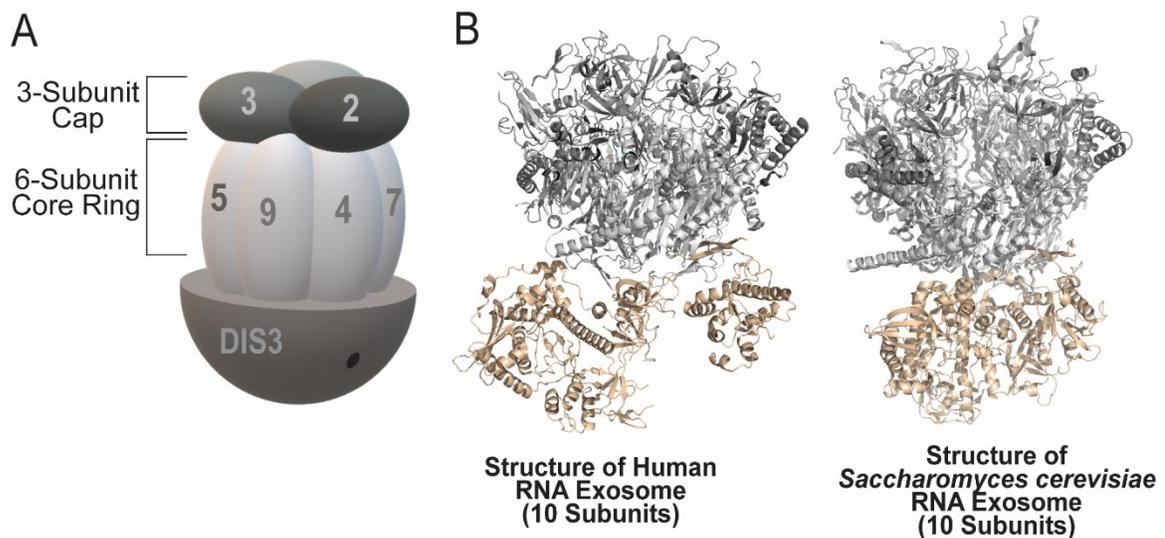


Figure 1. The RNA exosome is a multi-subunit complex that degrades/processes nearly every class of RNA. A) Diagram of the RNA exosome showing the three SI/KI cap subunits (EXOSC1/2/3), six PH-like domain core subunits (EXOSC4/5/6/7/8/9), and the catalytic DIS3 subunit. RNA is threaded through the cap and core to the catalytic DIS3 subunit. B) The RNA exosome is highly conserved in both structure and sequence. Shown are the solved structures of the human RNA exosome [PDB 6D6R [3]] and the *S. cerevisiae* RNA exosome [PDB 6FSZ [5]]

dynamics and the stability of the complex *in vivo*. These mechanistic questions are important in the context of understanding the consequences of pathogenic amino acid substitutions, such as those identified in RNA exosomopathies, on the function of the

complex. RNA exosomopathies are a diverse family of disease that have been linked to missense mutations in RNA exosome subunit genes [10-16]. Resulting amino acid substitutions have been identified in several cap and core subunits and are linked to neurodevelopmental disorders (EXOSC5 MIM *606492), pontocerebellar hypoplasia (EXOSC3, MIM #614678; EXOSC8, MIM #616081: and EXOSC9, MIM #618065) and a novel syndrome called SHRF (Short stature, Hearing loss, Retinitis pigmentosa and Faicies) (EXOSC2, MIM #617763). These pathogenic amino acid substitutions occur in highly conserved regions of the subunits.

Given the diverse pathologies and tissue-specificity of RNA exosomopathy diseases, a prediction is that that these pathogenic amino acid changes result in distinct *in vivo* consequences that impact the function of the RNA exosome differently. One mechanism these pathogenic amino acid substitutions could be resulting in functional defects of the complex could be by disrupting the integrity of the complex, perhaps shifting the complex's rate of association/dissociation within cells. We can hypothesize that impacting the dynamics of RNA exosome assembly/disassembly could result in an overall decrease in formed, stable complexes at a given time, resulting in some severe consequences in cells that may need regulate, temporal gene expression such as neurons.

Here we present an experimental design and preliminary data aimed at answering if different RNA exosome pathogenic amino acid substitutions are impacting the dynamics of the complex. We utilized an *S. cerevisiae* model of a singular RNA exosomopathy mutation, *EXOSC-G198D*. The *S. cerevisiae* model of this *EXOSC2* mutation has a corresponding missense mutation in the orthologous gene *RRP4*, generating *rrp4-G226D*, and encodes for a variant of the cap subunit Rrp that harbors a glycine to aspartic acid substitution (An *in vivo* functional study characterizing this mutant model is summarized in Chapter II). Previous work has shown that the Rrp4 G226D variant doesn't have any significant decrease in steady state

protein level compared to wild-type Rrp4 and can associate with the complex [17]. However, if a wild-type Rrp4 is present, the mutant Rrp4 G226D is outcompeted from the RNA exosome and does not associate with other complex members [17]. These data suggest that the Rrp4 G226D variant has some biochemical effects on the integrity of the complex and, therefore, serves as a good model to initially design and implement the following experiments.

A1.2 Experimental Design

The design of this experiment is outline in Figure 2. The general concept is to utilize an inducible construct that expresses wild-type Rrp4 tagged with a human influenza hemagglutinin (HA) epitope. This inducible construct is then transformed into cells that are only expressing the Rrp4 G226D variant tagged with a Myc epitope. Within the background of these cells would be another Rrp subunit that is TAP tagged (we utilized a strain that has an integrated TAP tag on the C-terminal of one of the core subunit genes). Expression of the Rrp4-HA variant is induced, cell samples are taken over a timecourse and protein lysate is extracted. The TAP-tagged Rrp subunit is isolated from each time point through biochemical immunoprecipitation We then can assess the associated levels of either Rrp4 G226D-Myc or Rrp4-HA to quantify an exchange rate which would reflect if assembly/disassembly of the complex is impacted by the pathogenic amino acid substitution.

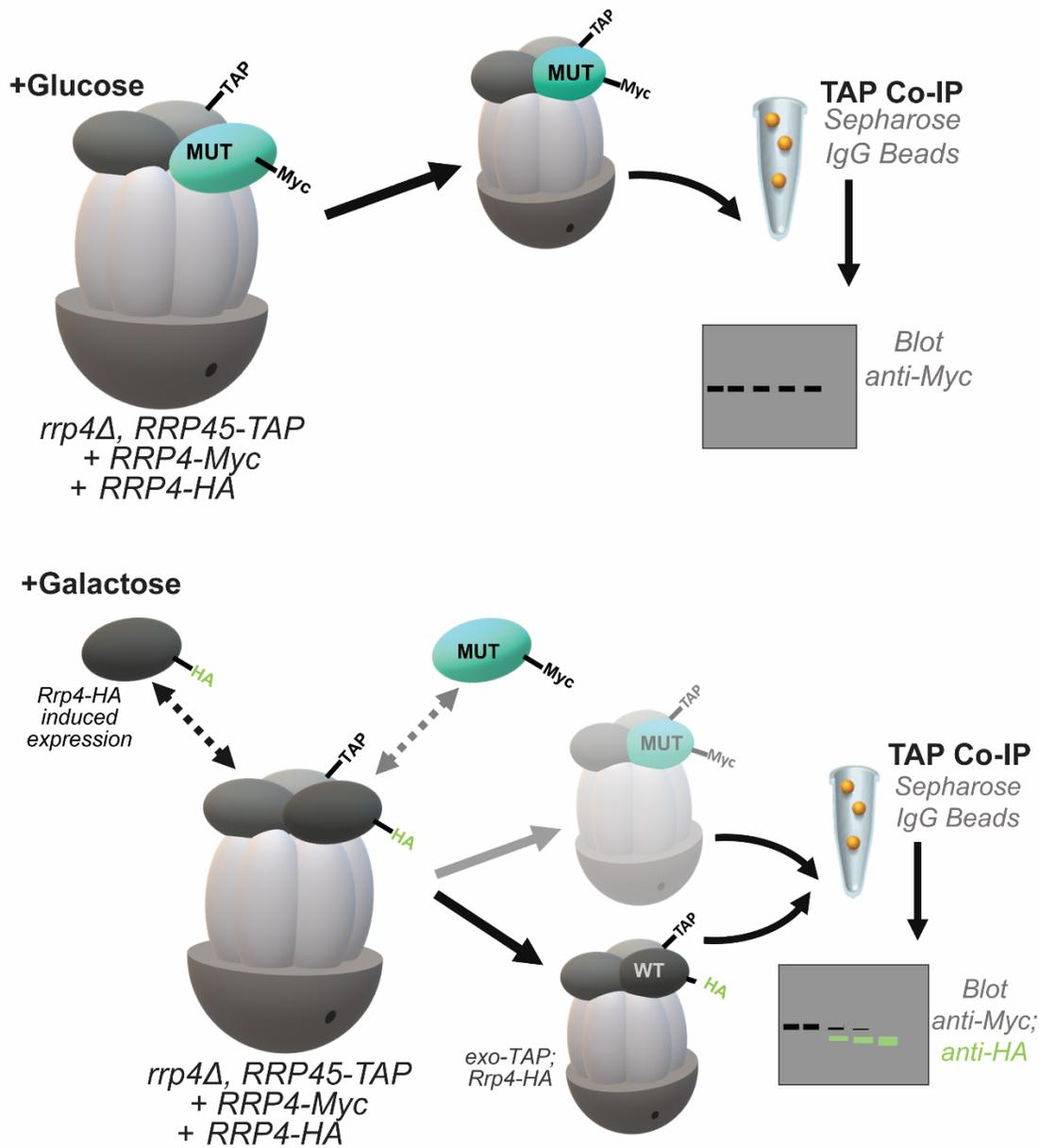


Figure 2. Experimental design to assess exchange of Myc tagged Rrp4 variant with HA tagged Rrp4.

Overall this experimental design can be divided into three steps:

- 1) Generating an inducible construct expressing wild-type Rrp variant with HA tag
- 2) Verifying expression of generated inducible construct

3) Performing the time point assay to assess exchange rate between the Myc tagged versus the HA tagged Rrp variant.

In the following text, we will present preliminary data from these three steps of the experimental design outlined in Figure 2.

A1.3 Step I: Generating Inducible Constructs Expressing wild-type Rrp variant with HA tag

The inducible plasmids were generated using the YCpGal1 plasmid (pAC16 in the Corbett lab; plasmid map available in Supplemental Figure 1) and NEBuilder HiFi DNA Assembly cloning techniques (NEB). Oligonucleotide primers were designed to amplify the wild-type *RRP* locus in accordance with overhangs that correspond to the YCPGal1 plasmid in accordance to NEBuilder HiFi standards. Additionally, oligonucleotides contained the coding sequence for a 2x HA epitope (*taccatacagatgttccagattacgct-taccatacagatgttccagattacgct*). The oligonucleotides used to generate a GAL-inducible plasmid expressing Rrp4-HA are summarized below:

<i>Generating Inducible plasmid expressing Rrp4 with 2xHA tag</i>		Tm	GC%
FWD (XbaI overhang)	ggatccactagttctagaATGTCCGAAGTTAT CACAATTACC	55C	38%
REV (3' end of <i>RRP4</i> CDS with HA epitope overhang)	agcgtaatctggaacatcgatgggtaGTTGCCGTT ACCTCTCATTT	54C	45%

FWD (2xHA epitope overhang with <i>RRP4</i> 3'UTR)	taccatacagatgtccagattacgcttaccatacagatgttc cagattacgctTAA TAGGAATACAAAAGC CGCTG	55C	39%
Rev (3'-5')	GTCTTGCACCTGCATCAGAAgtcgacctg caggcatgc		
REV (Sall Overhang)(5'-3')	gcatgcctgcaggctcgacTTCTGATGCAGGTG CAAGAC	57C	50%

Orange text corresponds to overhangs that overlap the added HA sequences.

Blue text corresponds to overhangs that correspond to the YCPGal1 plasmid.

As dictated by NEBuilder HiFi protocol, the YCpGal1 plasmid was digested before assembly. We digested the plasmid with XbaI and Sall, gel purified the products, and performed NEBuilder HiFi Assembly with the purified digested YCpGal1 and purified PCR products amplified with the above primers in accordance to NEBuilder HiFi protocol (NEB). Correctly assembled GAL-inducible plasmids expressing Rrp4-HA were isolated and verified by sequencing. Oligonucleotides were designed to generate GAL inducible plasmids that express other wild-type Rrp variants tagged with HA epitopes. All oligonucleotides are summarized in Table 1.

A1.4 Step II: Verifying expression of generated inducible construct

Once we verified the sequence of our GAL-inducible Rrp4-HA construct, we transformed yeast cells expressing the Rrp4-G226D variant as the sole copy of the cap subunit tagged with a 2x-Myc epitope (*rrp4-G226D-Myc*). We also transformed cells expressing a wild-type Rrp4-Myc variant (*RRP4-Myc*) with our GAL-inducible construct to use as a comparative control. Cells were then grown overnight to saturation in selective media containing 2% glucose. The next day, we induced Rrp4-HA expression using the following steps:

A) Spin down overnight cultures and wash yeast cells (3000 rpm x 3 minutes). Remove supernatant.

B) Resuspend pelleted cells in 1ml of water. Spin down and wash cells again (13,000 rpm x 1min). Remove supernatant.

C) Resuspend cells in fresh 1ml of water. At this point, take concentration of the washed, overnight cell sample. Using the measured concentration, calculate the volume for 5ml culture with an an OD of 0.2 using the following equation:

$$\text{Volume desired} = 5\text{ml} \times 0.2 \text{ OD} / (\text{OD of washed overnight cell sample})$$

D) Dilute the calculated volume of washed, overnight yeast cultures into 5mls of either selective media containing 2% glucose or selective media containing 2% galactose.

E) Grow cells at 30°C. After 2 hours, repeat steps A-B to spin down and wash cells.

Remove the supernatant after the final wash and freeze cell pellets.

Protein lysate was extracted from frozen cell pellets using RIPA buffer and standard lysis protocols of the Corbett lab (methods available in Chapter II and III). Levels of the Myc tagged and HA tagged Rrp4 variants were assessed by western blotting techniques. As shown in Figure 3, we see robust expression of the Rrp4-HA variant in both the *RRP4-Myc* and *rrp4-G226D-Myc* cells. Additionally, we can not detect any Rrp4-HA expression in either cells that were grown in 2% glucose, suggesting that our GAL-Rrp4-HA construct does not have any measurable leaky expression within this 2 hour induction window.

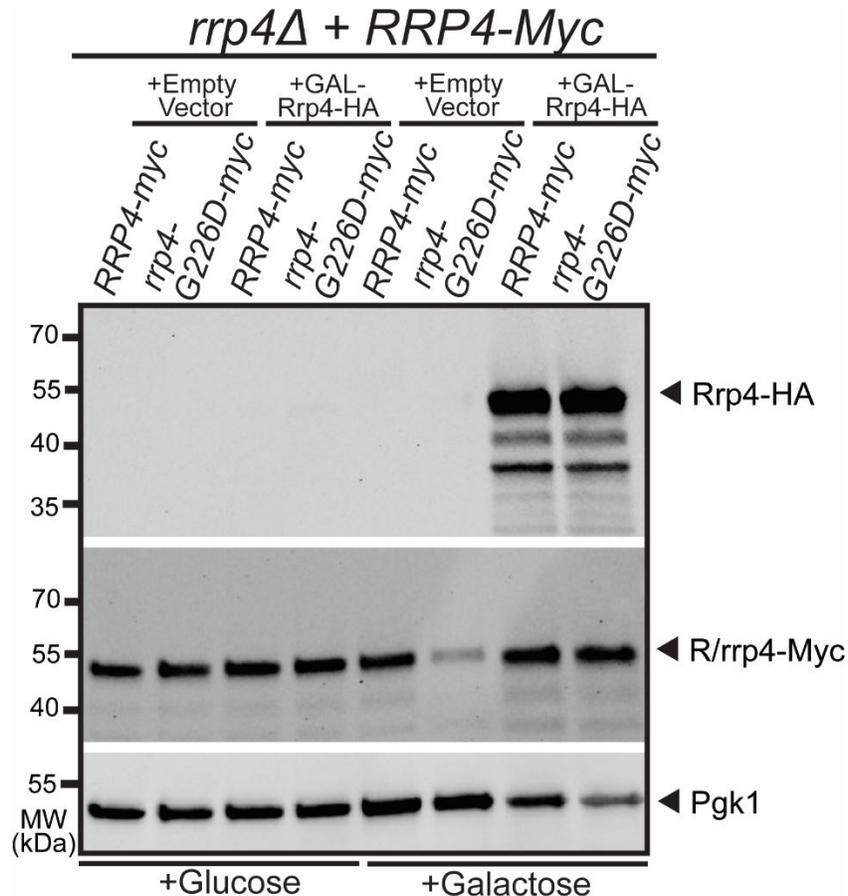


Figure 3. Rrp4-HA can be induced by robustly induced in cells expressing Rrp4-Myc or Rrp4 G226D-Myc as the sole copy of the cap subunit. *RRP4-myc* and *rrp4-G226D-Myc* cells were transformed with sequence verified Gal inducible Rrp4-HA constructs. Cells were also transformed with an empty vector to be used as control. Samples were incubated in 2% glucose or 2% galactose for 2 hours. Protein lysate was extracted, and sample concentrations were measure by a BCA assay. Rrp4-HA was detected by an anti-HA antibody; Rrp4-Myc and Rrp4 G226D-Myc were detected with an anti-Myc antibody. Pgk1 was blotted for and detected as a loading control.

A1.5 Step II:: Timecourse assay to assess exchange rate between the Myc tagged versus the HA tagged Rrp variant.

The goal of this experimental design is to perform a timecourse assay that would allow us to distinguish differences in Rrp4-HA association versus Rrp4-G226D Myc association with the RNA exosome complex. As shown in Figure 3, we can confirm that our GAL-Rrp4 HA construct expresses the wild-type Rrp4-HA

variant quite robustly with a two hour incubation. Therefore we next tried performing a timecourse assay to assess association of the Rrp4 HA variant in the complex over a two hour incubation.

For this assay, we utilized cells expressing the wild-type Rrp4 Myc variant as well as a TAP-tagged Rrp45 subunit (*RRP45-TAP*, *RRP4-Myc*). Rrp45 is a core subunit of the yeast RNA exosome and the TAP tag is integrated into the endogenous *RRP45* locus. We can then isolate Rrp45-TAP tagged subunit using IgG Sepharose beads and blot for the associated complex members. This methodology is described in detail in Chapter II and Chapter III.

Using overnight cultures of *RRP45-TAP*, *RRP4-Myc* cells transformed with the GAL-Rrp4 HA construct, we performed the following induction protocol:

A) Spin down overnight cultures and wash yeast cells (3000 rpm x 3 minutes). Remove supernatant.

B) Resuspend pelleted cells in 1ml of water. Spin down and wash cells again (13,000 rpm x 1min). Remove supernatant.

C) Resuspend cells in fresh 1ml of water. At this point, take concentration of the washed, overnight cell sample. Using the measured concentration, calculate the volume for 15ml culture with an OD of 0.2 using the following equation:

$$\text{Volume desired} = 15\text{ml} \times 0.2 \text{ OD} / (\text{OD of washed overnight cell sample})$$

D) Dilute the calculated volume of washed, overnight yeast cultures into 15mls of selective media containing 2% glucose. Additionally, dilute the calculated volume of cells into eight 5ml cultures of selective media containing 2% galactose. These eight samples will correspond to different time points.

E) Place cultures at 30°C. Immediately collect one of the eight 15ml + Galactose cultures and wash the cells as directed in steps A)-B). After final wash, remove the supernatant and freeze the cell pellet. This sample is considered the “0 minute” time point.

F) Let the other samples continue to grow at 30°C for two hours, collecting, washing, and freezing the cells from the remaining eight 15ml+Galactose cultures after 30 minutes, 45 minutes, 60 minutes, 75 minutes, and 90 minutes. At the end of the two hour timecourse, collect the final remaining 15ml+Galactose culture as well as the 15ml+Glucose culture. Spin down and wash the cells. Following the the final wash, remove the supernatant and freeze the cell pellets.

Protein lysate was extracted from frozen cell pellets using IPP150 buffer and lysis protocols as described in Chapter II and III. To isolate Rrp45-TAP, protein lysate was incubated with IgG Sepharose beads overnight at 4°C. Following incubation, IgG beads were washed and boiled. Bound protein was detected through western blotting techniques. Preliminary data shown in Figure 4A demonstrates that we can co-immunoprecipitate Rrp4-HA at differing levels across the timecourse assay. As we predict based on the experimental design, we see an increase in associated Rrp4-HA the longer samples were induced in 2% galactose. However this increase in co-immunoprecipitated Rrp4-HA over the timecourse does not appear linear as we anticipated. Rather, there appears to be a large jump in co-immunoprecipitated Rrp4-HA at 45 minutes induction compared to 30 minutes induction. This high level of co-immunoprecipitated Rrp4-HA detected in the 45 minute time point seems fairly constant between the 60 minute, 75 minute and 90 minute time points. This nonlinear increase in Rrp4-HA association with Rrp45-TAP may somewhat be explained by the nonlinear increase of induced Rrp4-HA as demonstrated in the input lanes in Figure 4A. After 45 minutes there is no distinguishable difference in Rrp4-HA input levels in the following time points,

suggesting that expression is outside of the linear range of detection in the later points of the timecourse assay.

As we want to be able to distinguish a difference in associated levels of the HA tagged Rrp4 variant using our experimental set up, we determine that we need to adjust the induction timecourse to capture a linear increase in induced Rrp4 HA levels. Therefore we decided to assess levels of Rrp4 HA expression over a shorter timecourse. Using overnight cultures of *RRP4-Myc* cells transformed with the GAL-Rrp4 HA construct, we performed the same induction protocol as described above however reducing the overall time to one hour and collecting samples at 30 minutes, 35 minutes, 40 minutes, 45 minutes, 50 minutes, 55 minutes, and 60 minutes.. Protein lysate was extracted from frozen cell pellets using RIPA buffer and standard lysis protocols (methods available in Chapter II and III) and levels of induced HA tagged Rrp4 variant were assessed by western blotting techniques. From the shorter timecourse experiment, we see a more linear increase in Rrp4 HA protein levels (Figure 4B). Additionally, the Rrp4 Myc protein levels are relatively even across these time points. These data suggest that can capture a measurable difference in Rrp4-HA over the shorter incubation times. Furthermore, this result can inform us of how to perform the timecourse assay originally designed and detailed in figure 2.

given that Gene Ontology (GO) analyses of differentially expressed genes in some of our RNA exosomopathy mutant models show biosynthetic and metabolic processes particularly impacted by the presence of the pathogenic amino acid substitutions.

Another important consideration for the experimental design is the turnover of the RNA exosome complex. We do not know how stable the RNA exosome complex is in the cell. Right now, we are assuming that approximately 1-2 hours is a long enough time for the induced Rrp4-HA to get into a complex. However depending on how often the RNA exosome complex disassociates, this time may not be long enough for the Rrp4-HA to properly incorporate into a new complex. Additionally, we do not have any way of knowing if inducing expression of an Rrp protein isn't just pushing the cell towards building more RNA exosome complexes. Therefore the association we see between Rrp4-HA and Rrp45-TAP, as shown in figure 4A, may not be that the induced variant is outcompeting the mutant variant, but rather the induced variant is increasing the number of new RNA exosome complexes. If the latter is the case, these data could still be interesting if there are differences between *rrp4-G226D Myc* or *RRP4-Myc* cells expressing Rrp4-HA as that would signify the Rrp4 G226D variant does impact the ability of complex formation. However, the calculated rate wouldn't entirely reflect an "exchange" or subunits as was originally devised. One idea to overcome some of these biases would be to induce expression for a set amount of time then take time points much later, giving more time for the potential dissociation/association of RNA exosome complexes within the cells and, thus, potential exchange of the Rrp4-HA variant with the Myc tagged variants. In order to perform this experiment it would be pertinent to first assess half life of the Rrp4-HA in order to know the length of time needed to capture an exchange of the subunits but not a degradation of the induced variant.

In conclusion, there are still many directions to take this experiment. Ultimately, assessing these mechanistic questions can lead to a deeper understanding of the *in vivo* functional consequences these RNA exosomopathy pathogenic amino acid substitutions have on the RNA exosome complex. The different functional consequences resulting on the structure of the complex may also help us understand the diversity of disease pathologies seen in RNA exosomopathies. In addition, the types of experiments and tests outline

here will expand our knowledge of the basic biology of the RNA exosome as they will shed light on how the complex assembles/disassembles.

Table 1. Plasmids and Oligonucleotides

Identifier	Description
pAC16	pGAL; URA3; CEN; ampR
AC9617	Fwd XbaI overhang GAL1 Rrp4
AC9618	Rev Native 3' Rrp4 with HA tag overhang
AC9619	Fwd 2X HA overhang Rrp4
AC9620	Rev SalI overhang GAL1 Rrp4
AC9621	Fwd GAL1 inducible exo subunits validation seq oligos
AC9622	Rev GAL1 inducible exo subunits validation seq oligos
AC9796	GAL induced Rrp40 2x HA Tag Building FWD 1
AC9797	GAL induced Rrp40 2x HA Tag Building REV1
AC9798	GAL induced Rrp40 2x HA Tag Building FWD 2
AC9799	GAL induced Rrp40 2x HA Tag Building REV 2
AC9800	GAL induced Rrp46 2x HA Tag Building FWD 1
AC9801	GAL induced Rrp46 2x HA Tag Building REV1
AC9802	GAL induced Rrp46 2x HA Tag Building FWD 2
AC9803	GAL induced Rrp46 2x HA Tag Building REV 2
AC9804	GAL induced Rrp45 2x HA Tag Building FWD 1

AC9805	GAL induced Rrp45 2x HA Tag Building REV1
AC9806	GAL induced Rrp45 2x HA Tag Building FWD 2
AC9807	GAL induced Rrp45 2x HA Tag Building REV 2
AC9808	GAL induced Rrp43 2x HA Tag Building FWD 1
AC9809	GAL induced Rrp43 2x HA Tag Building REV1
AC9810	GAL induced Rrp43 2x HA Tag Building FWD 2
AC9811	GAL induced Rrp43 2x HA Tag Building REV 2

A1.7 References

1. LaCava, J. and S. Vanacova, *Correction to: The Eukaryotic RNA Exosome*. *Methods Mol Biol*, 2020. **2062**: p. C1-C4.
2. Morton, D.J., et al., *The RNA exosome and RNA exosome-linked disease*. *Rna*, 2018. **24**(2): p. 127-142.
3. Weick, E.-M., et al., *Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex*. *Cell*, 2018. **173**(7): p. 1663-1677.e21.
4. Mitchell, P., et al., *The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases*. *Cell*, 1997. **91**(4): p. 457-466.
5. Schuller, J.M., et al., *Structure of the nuclear exosome captured on a maturing preribosome*. *Science*, 2018. **360**(6385): p. 219-222.
6. Liu, J.-J., et al., *CryoEM structure of yeast cytoplasmic exosome complex*. *Cell Research*, 2016. **26**(7): p. 822-837.
7. Wasmuth, E.V., K. Januszyk, and C.D. Lima, *Structure of an Rrp6–RNA exosome complex bound to poly(A) RNA*. *Nature*, 2014. **511**(7510): p. 435-439.
8. Makino, D.L., et al., *RNA degradation paths in a 12-subunit nuclear exosome complex*. *Nature*, 2015. **524**(7563): p. 54-58.
9. Falk, S., et al., *Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase*. *Cell Reports*, 2017. **20**(10): p. 2279-2286.
10. Biancheri, R., et al., *EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement*. *J Neurol*, 2013. **260**(7): p. 1866-70.
11. Boczonadi, V., et al., *EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia*. *Nature Communications*, 2014. **5**.
12. Burns, D.T., et al., *Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy*. *Am J Hum Genet*, 2018. **102**(5): p. 858-873.
13. Di Donato, N., et al., *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*. *Journal of Medical Genetics*, 2016. **53**(6): p. 419-425.
14. Schottmann, G., et al., *Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia*. Vol. 37. 2017.
15. Slavotinek, A., et al., *Biallelic variants in the RNA exosome gene EXOSC5 are associated with developmental delays, short stature, cerebellar hypoplasia and motor weakness*. *Hum Mol Genet*, 2020. **29**(13): p. 2218-2239.

16. Wan, J., et al., *Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration*. *Nature Genetics*, 2012. **44**(6): p. 704-U134.
17. Sterrett, M.C., et al., *A budding yeast model for human disease mutations in the EXOSC2 cap subunit of the RNA exosome complex*. *Rna*, 2021. **27**(9): p. 1046-1067.

Appendix III. Amplifying Growth Mindsets With PCR: Implementing growth mindset interventions in introductory biology lab to increase resiliency in undergraduate STEM students.

The following was a “Teaching As Research” project. This project was designed and implemented as part of my experience as an Emory Advanced Graduate Teaching Fellow during the academic year 2020-2021.

A3.1 Introduction

The “world of science” is messy. Scientific discovery and research are full of failures and setbacks, requiring constant reevaluation, critical analysis, and creativity to achieve success. Introduction to this world tends to be presented in a prototypical format, with students following step-by-step instructions with guaranteed results, requiring little to no creative problem solving or analysis [1, 2]. In addition to what students experience in the classroom, media portrayals of the “white lab coat” scientists show instant scientific success—case in point, any medical drama or crime scene investigation where molecular techniques are simple scene fillers with no apparent error or failure. Previous research suggests that these two influences, coupled with other societal experiences, cause our students to enter college ill-equipped to view failures and challenges as learning experiences [3, 4]. In addition to our students entering college without the framework to cope with failure, we have seen a persistence issue within STEM (science, technology, engineering, and mathematics) fields in college. Bachelor’s degrees awarded in the STEM fields have not increased substantially over the past 40 years in the United States as reported by the National Science Foundation [5], grossly under pacing the significant rise in total awarded degrees. This is worrying trend given the more than 3 million job openings in STEM that are predicted with the retirement of baby-boomer generation [6] as well as the growing societal need of advanced scientific discoveries exemplified by the recent coronavirus pandemic and growing environmental crisis. Previous framework to address this STEM workforce disparity has focused on increasing rigor and performance, however recent studies suggest that several intrinsic factors, such as mindset and motivation, play a larger factor in curtailing STEM dropout among other non-cognitive factors. Navigating failure and persisting through difficulties are cited as some of the

most important dispositions distinguishing scientists [4, 7-9], thus in order to advance STEM we must also focus on composing challenge-engaging individuals who have the ability to persevere through failure.

The question to ask then is how and when do we develop dispositions in our students that will have them embrace challenge and failure? To address the former, Henry et al.’s 2019 article “FAIL is not a Four Letter Word: A theoretical Framework for Exploring Undergraduate Students’ approaches to academic challenge and response to failure in STEM learning environments” generated a failure mindset coping model that integrated previous work to describe how STEM

students engage with challenges and respond to failure [1]. Henry et al. illustrate how five constructs—mindset, goal orientation, fear of failure, attributions, and coping responses—interplay and influence a student’s ability to respond to failure (Figure 1). This framework is extensive, but limited in it excludes contextual, pedagogical and demographic factors and assumes that success is a desired outcome by students. Furthermore, the framework is based mainly on correlative work that includes no randomized, experimental studies. However, this framework does collectively draw on theory from a variety of fields and can serve as a guide to investigations, particularly those designing and testing

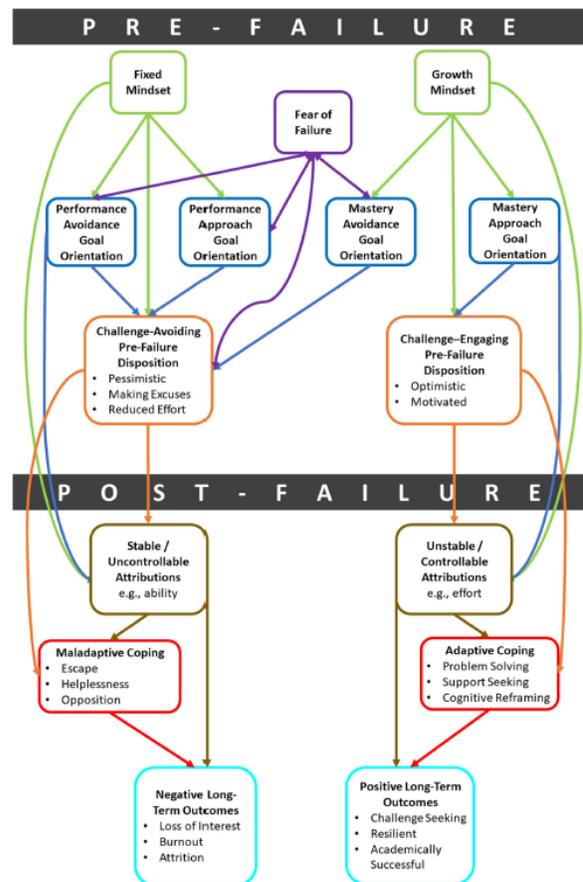


Figure 1. The failure mindset coping model. All relationships (solid arrows) represent predicted relationships between constructs in undergraduate STEM contexts but are supported by previous work outside of STEM contexts [1].

interventions aimed at assessing student dispositions pre and postfailure. While all five constructs of Henry et al.'s model are of interest for future study, the construct of mindset likely affects all subsequent factors, either directly or indirectly, as detailed by the many connecting arrows in figure 1. Additionally, Henry et al.'s model postulates that mindset could play a key role in influencing STEM undergraduate engagement with a challenge *before* the failure occurs, thus suggesting that creating interventions to shift student mindset could likely increase their ability to navigate *future* failures and gain the dispositions needed to become successful scientists.

A3.2 Defining “Mindset”

The term “Mindset” was first introduced by researcher Dr. Carol Dweck in 1999 [10] and gained worldwide fame in 2006 with the publication of her best seller *Mindset: The New Psychology of Success* [11]. Dr. Dweck repeatedly observed that individuals with similar skills and abilities experienced drastically different outcomes, with some achieving successes and others not, and those with less abilities or skills rising to great heights of some with “raw talent” through sheer perseverance and will power. This phenomena occurred across fields and domains, including classrooms, board rooms and sports fields [11]. Dr. Dweck concluded that success is less a result of one's abilities or skills but of one's beliefs about one's abilities or skills. Thus, she proposed two “mindsets”; a “fixed mindset” in which one believes that intelligence and skills are unchangeable, inherent traits, and a “grow mindset” in which one believes these qualities malleable and can grow over time and through effort. Since Dr. Dweck's groundbreaking theory, several studies have shown that our mindset has profound effects on our self-perception [12], goals for learning [13, 14], and approach to challenges [15]. In regard to an approach to challenges, studies show that individuals with a fixed mindset are more likely to avoid an academic challenge and see

learning opportunities only as ways to “prove” their talent or intelligence [16]. In contrast, individuals with a growth mindset seek out challenges and see learning opportunities as a change to improve [10, 12, 15, 16].

Mindset is also linked to other non-cognitive factors that affect how students approaches challenges and respond to failure. Studies have shown that students with a growth mindset generally have a reduced fear of failure, have goals oriented towards mastering the challenge, and adapt to failure with problem-focused coping strategies that allow for them to learn and grow from the challenge[14, 17-23]. Contrastingly, students with a fixed mindset have increased fear of failure, have goals oriented towards avoiding the challenge, and cope with failure through unproductive strategies such as venting, avoidance and distancing from the perceived problem[14, 18-20, 22-24]. The qualities found associated with a growth mindset in students are those that underlie the disposition of a successful scientist. As depicted in figure 1 and Henry et al.’s model, a growth mindset in STEM students pre failure influence responses post failure as well as sets the student up for productive response to any future challenges. Thus, by introducing interventions to push our STEM undergraduates towards a growth mindset will not only allow them to face failures and challenges productively, it may help in the development of the next generation of challenge-engaging perseverant scientists.

A3.3 Rationale and Approach

Therefore, in addressing our question we choose to frame the “how” through mindset theory, designing an intervention to increase growth mindset in our STEM students. The second part then is “when”. Introductory lab courses serve as a perfect opportunity to influence student mindset and, thus, their disposition towards one of a scientist. Introductory college courses serve

as a gateway into STEM fields, and student achievement in these courses is of utmost importance for academic motivation and retention in STEM majors [25]. Furthermore, introductory laboratory courses can serve as a gateway to the “messiness” of science. Introductory laboratory courses, particularly in course-based undergraduate research experiences (CUREs), can serve as authentic introductions to research, as students have hands on experience conducting scientific research and creating new knowledge [2]. Emory’s Introductory Biology Laboratory courses are an example of student-centered educational approach towards biology laboratory instruction, particularly the Bean Beetle Microbiome Project. The Bean Beetle microbiome Project is a National Science Foundation (NSF) funded initiative to develop bean beetles as a model system for undergraduate laboratories [26]. Through the Bean Beetle Microbiome Projects, students grow larvae on different foods, extract the microbiome, grow up the microbiome and identify bacterial species through morphology and genetic analysis. This research is conducted across several introductory biology courses and has been developed into a “beetle handbook” which includes three dozen lab modules developed by Emory faculty [27].

One module includes students performing a Polymerase Chain Reaction (PCR) on isolated bacterial DNA and sequencing the DNA to identify the bacterial species via BLASTn alignment searches [27]. This module is conducted every spring semester during Emory College’s Introductory Biology Laboratory Course II BIO142, mostly by first-year college students. This lab module is many students first experience with molecular techniques, such as pipetting enzymatic reagents in small volumes and running an agarose gel. This lab module is therefore riddled with “failure” and many times results in students having to problem solve and reassess in order to complete the lab requirements.

This lab therefore serves as a perfect, controlled environment to observe how a growth mindset intervention influences student's ability to navigate through failure and setbacks. As supported by previous work, increasing growth mindset will allow students to have mastery goal orientation and challenge-acceptance disposition, which may further increase their cognitive understanding of the biological principles underlying the lab module (ie PCR and gel electrophoresis). Furthermore, Henry et al.'s framework suggests that increasing student's growth mindsets in an introductory course could have long lasting impacts long-term outcomes in retaining STEM students. Therefore, I chose as my Teaching as Research (TAR) project to introduce a mindset intervention in the BIO142 Bean Beetle PCR module to assess student cognitive and non-cognitive abilities.

I hypothesize the following:

- 1) A growth mindset intervention, relative to control, will lead to stronger growth mindset scores in relation to research science.
- 2) A growth mindset intervention, relative to control, will increase student's cognitive abilities to solve higher order problems regarding PCR/gel electrophoresis techniques.
- 3) A growth mindset intervention, relative to control, will increase a student's resiliency (as measured through GRIT-S scale).
- 4) The growth mindset intervention will correlate with increased motivation in pursuing hands on, research-based science.

A3.4 Experimental Design

The mindset intervention I will implement is modeled after the interventions Burnette et al. implemented in their 2019 study [28]. Their novel mindset intervention incorporated elements

of previously published mindset interventions, including explanations of research related to growth mindset [29, 30], use of a role model to strengthen attitude change and motivation [31, 32], and a “saying is believing” writing exercise use in past interventions [33]. My mindset intervention will be slightly modified given the COVID-19 restrictions and digital medium. Therefore, all intervention materials will be given through asynchronous online learning platforms.

To quantify the influence of the designed mindset intervention, students will also be given pre intervention and post intervention assessments that will be administered through Qualtrics. In addition to a post intervention assessment, students will be given a cognitive “quiz” that has questions differing in Bloom’s taxonomy of learning. This quiz will have no grade weight and be given purely as an exercise during the synchronous lab sessions. The quiz will be administered through Qualtrics, with each question weighted differently to assess if the mindset interventions impacted students cognitive comprehension of the course material. Lastly, at the end of the spring semester, students will be administered a survey to investigate students motivation in continuing in STEM. Student scores will be matched to their mindset intervention groups to identify if there is any correlative relationship between a growth mindset and persistence in STEM.

Participants

The participants of this study will be spring 2021 BIO142 students (approximately 650 students). The BIO142 laboratory course is broken into 29 sessions led by different mentors. Students will be separated into experimental or control groups based on their registered laboratory session. Sessions were divided into “experimental” or “control” category by random:

Experimental	Control
Th 2:40pm - Mentor Malay	Th 8am - Mentor Maria

W 11:20am - Mentor Nourine	Th 2:40pm - Mentor Alex
W 2:40pm - Mentor Piper	F 2:40pm - Mentor Brooke
F 8am - Mentor Megan	F 8am - Mentor Ciavatta
F 11:20am - Mentor Matthew	W 11:20am - Mentor Piper
W 8am - Mentor Ariel	F 11:20am - Mentor Megan
W 2:40pm - Mentor Megan	W 11:20am - Mentor Maxwell
W 2:40pm - Mentor Annie	W 2:40pm - Mentor Ben Cai
Th 2:40pm - Mentor Brooke	Th 11:20am - Mentor Alice
F 6pm - Mentor Alexia Rodriguez	W 8am - Mentor Kane cooper
Th 6pm - Mentor Kane Cooper	Th 2:40pm - Mentor Dom
F 11:20am - Mentor Dom	W 11:20am - Mentor Lewis
W 6pm - Mentor Gillespie	W 2:40pm - Mentor Ariel
Th 11:20am - Mentor Ciavatta	
F 2:40pm - Mentor Alisha	Th 6pm - Mentor Lauren

Experimental Intervention—Three part online structure: As modeled by Burnette et al.’s 2019 study [28], the interventions will be three parts. The first two parts will be given prior to the BIO142 Bean Beetle PCR module first lab session, with the third final part given prior to the BIO142 Bean Beetle PCR module second lab session.

- 1) Students will be introduced to the research topic and benefits of a growth mindset by reading “You can grow your brain” [29, 30] (Appendix 1)

2) Students will participate in a short “saying is believing” writing exercise using an indirect framing approach as has been previously shown more affective in generating growth mindsets [29].

- a. Prompt: “Students often do a great job explaining ideas to their peers because they see the world in similar ways. We would like your help to explain the information you just read in a more personal way that students will be able to understand and apply it to their scientific research and the upcoming lab module”.

**(Adapted from Yaeger et al [29])

3) Students will watch short video that reiterates the message “you can develop your research skills and ability through hard work” through a role model. The role model will be a doctoral student or post-doctoral fellow who will give a tour of a molecular lab and show a PCR machine and gel electrophoresis box. Additionally, the role model will tell a brief story of how a PCR reaction or gel electrophoresis failed, but through effort and hard work they were able to resolve the problem and adapted to it. <https://youtu.be/uwazjOgVF7g>

Attention-matched control—Three part online structure: As modeled by Burnette et al.’s study [28], there will be an attention-matched control group. The aim is for these control interventions to be similar to the experimental in terms of length and style. As with the experimental intervention, the control intervention is three parts, with the first two parts given prior to the BIO142 Bean Beetle PCR module first lab session and the third given prior to the BIO142 Bean Beetle PCR module second lab session.

1) Students will read an interest article about Kary Mullis’ discovery of PCR.

(<https://www.mcgill.ca/oss/article/technology-history/man-who-photocopied-dna-and-also-saw-talking-fluorescent-raccoon>).

- 2) Students will participate in a short writing exercise, similar in length as the experimental intervention.
 - a. Prompt: “Students often do a great job explaining ideas to their peers because they see the world in similar ways. We would like your help to explain the information you just read in a more personal way that students will be able to understand Kary Mullis’ discovery?”
- 3) Students will watch a short video with the same role model as the experimental intervention, but without the reiteration of developing one’s skills and ability. Within the control video, the role model will solely give a tour of a molecular lab and show a PCR machine and gel electrophoresis box. <https://youtu.be/gnHSLPjnu0I>

Pre and Post Intervention Assessments

The assessments used in this project will be modeled after those used by Burnett et al. in their 2019 study [28]. In it, Burnette et al. adapted an established mindset measurement (the Growth Mindset Scale) [10] to the domain of computer science by replacing the word “intelligence” with “computer science”[28]. I will take the same approach, only adapting the Growth Mindset Scale to the domain of lab research skills, replacing the word “intelligence” with “science research” (Appendix 2). Similar to the original measurement, the higher score will represent a stronger orientation toward growth mindsets of lab research. Additionally, I will administer the short GRIT-S scale [34, 35] (Appendix 3).

The post intervention assessments will also include a cognitive skills assessment in the form of a 5 minute “quiz”. This quiz will consist of four questions, each testing a different level of Bloom’s Taxonomy of Learning. Question 1 will assess “Knowledge”, question 2 will assess “Comprehension”, question 3 will assess “Analysis” and question 4 will assess “Evaluation” (See

Appendix 4). The questions will be weighted more heavily based on the Bloom's level, and therefore a higher score will represent increased cognitive understanding of the material. I will ask lab mentors in both the control and experimental groups to administer this "quiz" through Qualtrics.

Post semester survey

For the post semester survey to assess student motivation and persistence in STEM, I will utilize the Science Motivation Questionnaire II (SMQII) developed by Glynn et al. in 2011 [36]. SMQII assesses five scales: intrinsic motivation, self-determination, self-efficacy, career motivation, and grade motivation. Each scale has five questions, totaling 25 questions designed to be answered in only 15 minutes (see Appendix 5). This survey will be given through Qualtrics and can be attached to the students end of term evaluations.

A3.5 Results and Limitations of study

From the study design, I received a total response rate of 40-46% (n for pre-intervention = 251, n for post intervention = 208). However, I only had a 9-24% response rate for paired responses (n control = 28, n experimental = 52). These paired responses indicated that the growth mindset intervention did not significantly increase BIO142 students' growth mindset scores (Figure 2A). Overall, BIO142 students had high growth mindset scores pre intervention, with the average for both the experimental and control groups ranging between 4.8-5.04. Intriguingly however I did observe a significant decrease in

the control groups post intervention growth mindset scores ($p < 0.05$). These data suggest that while I did not see an improvement on growth mindset surrounding the PCR lesson, the intervention may have protected a decline in student growth mindset over the time of the course.

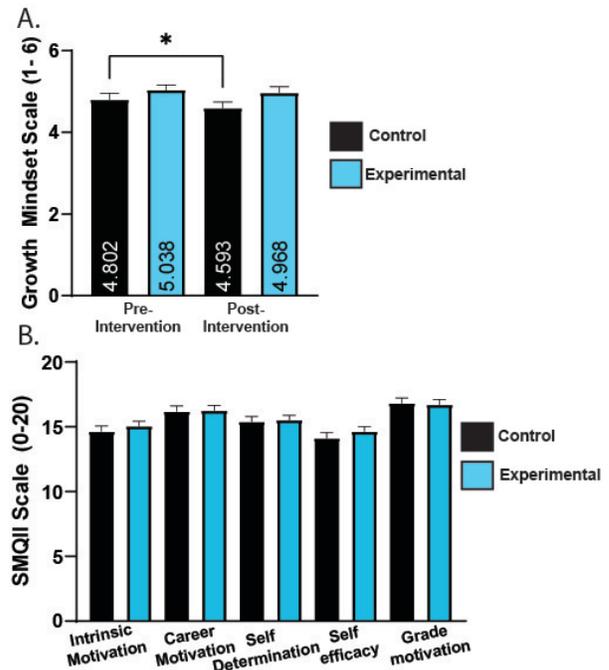


Figure 2. Pre and Post Intervention Mindset and SMQII results. A) The average paired growth mindset score responses for the experimental (blue) and control (black) groups are shown. Paired t-tests were performed between pre and post intervention groups. Asterisk * $p < 0.05$. B) The average paired SMQII score responses for the experimental (blue) and control (black) groups are shown.

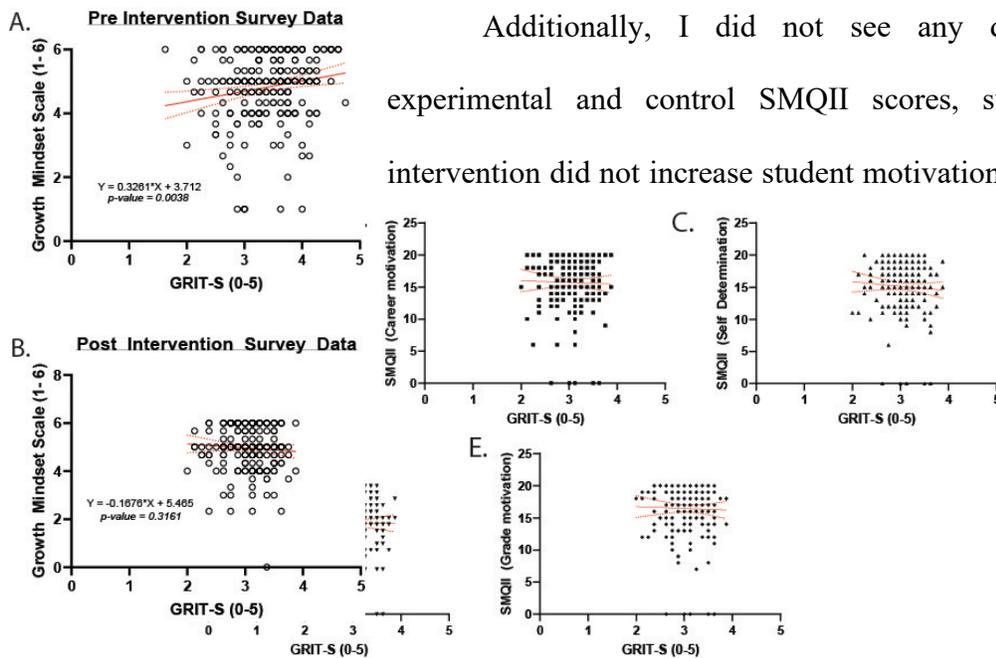


Figure 3. The relationship between resiliency and student motivation in STEM as measured by SMQII scales. Post intervention GRIT-S responses were graphed against student SMQII scores. Simple linear regression tests were performed (in red). The SMQII scales include A) Intrinsic motivation, B) Career motivation, C) Self determination, D) Self efficacy, and E) Grade motivation.

Additionally, I did not see any difference between experimental and control SMQII scores, suggesting that the intervention did not increase student motivation for science (Figure 2B). Given the similarity in SMQII scores between the two groups, I wanted to assess if student motivation was intrinsically linked to resiliency as measured by the GRIT-S scores. I performed linear regression analysis between the post intervention GRIT-S scores and the SMQII scores, breaking down SMQII scores into the five scales measured (intrinsic motivation, self-determination, self-efficacy, career motivation, and grade motivation). As shown in figure 3, I saw no linear correlation between the SMQII scales and GRIT-S scores. To further assess the interplay between student resiliency and a growth mindset, I performed simple linear regression analysis on total GRIT-S scores and total growth mindset scores. I see a positive correlation between GRIT-S scores and growth mindset scores pre intervention (Figure 4A), however that positive relationship is lost post intervention (Figure 4B). The loss of correlation between the GRIT-S scores and the growth mindset scores post intervention may in part be due to the decrease in GRIT-S scores I observed (Figure 5). When I assessed the GRIT-S scores between my experimental and control groups, I saw a shared decline between the pre and post intervention datasets (Figure 5A). From the paired responses, I saw a

significant decrease in GRIT-S scores between both the experimental and control groups of 0.352 and 0.465, respectively. I then assessed the overall GRIT-S scores from the total responses I received and saw a significant decrease between the pre and post intervention datasets (Figure 5B). Overall, these data suggest that BIO142 students may have experienced a decline in resiliency throughout the semester that may be impacting my results.

In part, I believe this is due to the effects of the COVID-19 pandemic. During this semester, BIO142 students were remote learning and did not take part in any hands on activities during the

Figure 4. A positive relationship between resiliency and growth mindset pre intervention. A) Pre and B) post intervention GRIT-S responses were graphed against student growth mindset scores. Simple linear regression tests were performed (in red).

lab course. Additionally, students had been virtually learning for the entire academic year, and most likely were experiencing online learning fatigue which I believe ultimately is reflected in these negative GRIT-S scores. In addition to the confounding variable of remote learning during a pandemic, my study also had several limitations that impacted my conclusions. I

had only received a 9-24% paired response rate, making my sample size small for the comparative analysis. Additionally, due to the nature of the study design, I only performed this intervention for a short period of time. The Burnette et al. study [28] designed an experiment over the course of a full semester with multiple reiterative interventions. Due to time constraints, I was only able to implement the intervention for one module of the BIO142 course which meant students were exposed to material over a three week time period with no reiteration and repetition. Lastly, I did

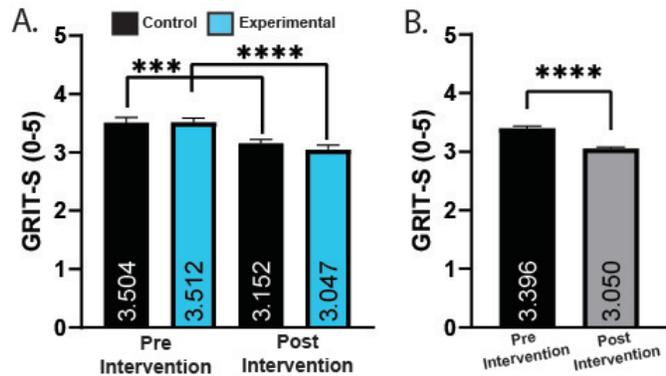


Figure 5. Decrease in GRIT-S scores in BIO142 students. A) The average GRIT-S scores of the paired responses are graphed, with the control group in black and the experimental group in blue. B) Average GRIT-S scores of total responses pre and post intervention are graphed. Paired t-tests were performed. Asterisks ***p<0.001 and ****p<0.0001.

not collect any demographic data on the experimental and control participant students. Therefore my data may also have some confounding variables based on students' lived experiences. This in part is also a limitation of the Henry et al [1] framework that was the basis for this work, as this framework excludes demographic and contextual factors.

A3.5 Discussion

Overall, this study had many limitations and was inconclusive in predicting if there was an increase in growth mindset within BIO142 students after a designed intervention. Additionally, the growth mindset intervention did not correlate with increased motivation in pursuing hands on, research-based science in BIO142 students. It should be noted that the control group of BIO142 students had fairly high SMQII scores (Figure 2B). This suggests that the student population in BIO142 may already have increased motivation to stay within the sciences and that the SMQII scale was not an appropriate measurement to assess if the intervention surrounding the PCR methodology impacted student views on failure at the bench. My results did however suggest that BIO142 student resiliency measured by the GRIT-S scores decreased significantly over the course of the growth mindset intervention study. In part this could be due to the ongoing COVID-19 pandemic and remote learning students were undergoing at the time. I believe that this decreased resiliency in the wake of virtual learning raises questions about how to build resiliency in online platforms. For example, how can we engage in challenging STEM research online? Also how can we really evaluate resiliency in STEM students through online platforms?

This leads to my next point in thinking about “resiliency” within our STEM students. While I believe exposure to hands on bench work is rife with potential build students' mindsets, I also recognize that we must approach such topics equitably and humanistically. Within this study I utilized the GRIT-S score generated by Dr. Angela Duckworth , whose research and assertions on

gritiness of students has come under criticism as being “deficient thinking” that can disproportionately affect marginalized students. The idea that success is solely based on one’s “grit” and “mindset” ignores the very real social environment that can profoundly impact our students. Furthermore, a focus on grit takes an impoverished view of human motivation. To some effect, I also believe this is reflected in my data showing little correlation between motivation in STEM and grit, as measured through the SMQII and GRIT-S scores (Figure 3). To quote Dr. Jal Mehta, a professor at the Harvard Graduate School of Education, “...most people do not persevere at things because they are good at persevering, they persevere because they find things that are worth investing in...”. Dr. Mehta is speaking to motivation in finding authentic purpose and engaging in one’s passions. Therefore in thinking about future steps with this type of research, I would shift the focus to building “passion” and “purpose”. This approach would not only be more equitable but more indicative of future success for our students, be it in STEM or in other fields.

A3.6 Appendices

Appendix 1 “You can Grow Your Brain” Reading [29, 30]

You Can Grow Your Brain

New Research Shows the Brain Can Be Developed Like a Muscle

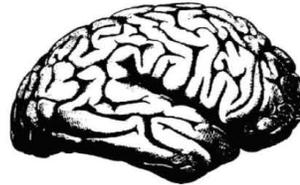
Many people think of the brain as a mystery. We don't often think about what intelligence is or how it works. And when you do think about what intelligence is, you might think that a person is born either smart, average, or dumb—either a “math person” or not—and stays that way for life.

But new research shows that the brain is more like a muscle—it changes and gets stronger when you use it. Scientists have been able to show just how the brain grows and gets stronger when you learn.

Everyone knows that when you lift weights, your muscles get bigger and you get stronger. A person who can't lift 20 pounds when they start exercising can get strong enough to lift 100 pounds after working out for a long time.

That's because muscles become larger and stronger with exercise. And when you stop exercising, the muscles shrink and you get weaker. That's why people say “Use it or lose it!”

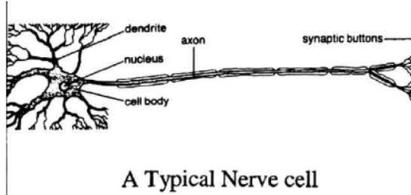
But most people don't know that when they practice and learn new things, parts of their brain change and get larger, a lot like the muscles do. This is true even for adults. So it's not true that some people are stuck being “not smart” or “not math people.” You can improve your abilities a lot, as long as you practice and use good strategies.



A Section of the Cerebrum nerve fibers (white matter)

Inside the outside layer of the brain—called the cortex—are billions of tiny nerve cells, called neurons. The nerve cells have branches connecting them to other cells in a complicated network. Communication between these brain cells is what allows us to think and solve problems.

When you learn new things, these tiny connections in the brain actually multiply and get stronger. The more you challenge your mind to learn, the more your brain cells grow.



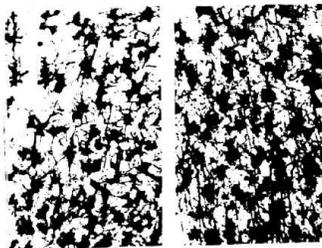
A Typical Nerve cell

Then, things that you once found very hard or even impossible to do—like speaking a foreign language or doing algebra—become easier. The result is a stronger, smarter brain.

How Do We Know That The Brain Can Grow Stronger?

Scientists started thinking the human brain could develop and change when they studied adult animals' brains. They found that animals who lived in a challenging environment, with other animals and toys to play with, were different from animals who lived alone in bare cages.

While the animals who lived alone just ate and slept all the time, the ones who lived with different toys and other animals were always active. They spent a lot of time figuring out how to use the toys and how to get along with other animals.



Nerves in brain of animal living in bare cage.

Brain of animal living with other animals and toys.

These animals had more connections between the nerve cells in their brains. The connections were bigger and stronger, too. In fact, their whole brains were about 10% heavier than the brains of the animals who lived alone without toys.

The adult animals who were exercising their brains by playing with toys and each other were also “smarter” –they were better at solving problems and learning new things.

Can Adults Grow Their Brains?

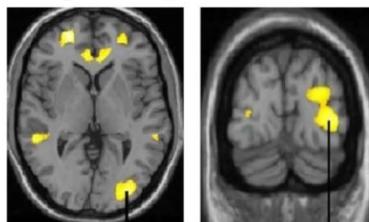
Scientists have recently shown that adults can grow the parts of their brains that control their abilities—like the ability to do math or even to juggle.

In one study, scientists found a group of adults who were not jugglers. They taught half how to practice juggling in the right way. These people practiced

for a long time and got much better at juggling. The other half didn't practice, and didn't get better.

Next, the scientists used a brain scanner to compare the brains of the two groups of people. They found that the people who learned how to juggle actually grew the parts of their brains that control juggling skills—the visual and motor areas. Their brains had changed, so they actually had more ability.

This was surprising because these people said before the study that they couldn't juggle—just like some people say they're "not good at math." But when they learned good strategies for practicing and kept trying, they actually learned and grew their brains.



In Yellow: Parts of the brain that grew when adults learned to juggle

doi:10.1371/journal.pone.0002669.g001

This can happen because learning causes permanent changes in the brain. The jugglers' brain cells get larger and grow new connections between them. These new, stronger connections make the juggler's brain stronger and smarter, just like a weightlifter's toned muscles.

***A Formula For Growing Your "Math Brain":
Effort + Good Strategies + Help From Others***

Scientists have also found that learning to juggle is a lot like getting better at math. When people learn and practice new ways of doing algebra or statistics, it can grow their brains—even if they haven't done well in math in the past.

Strengthening the "math" part of your brains usually happens when you try hard on challenging math problems. But it's not just about effort. You also need to learn skills that let you use your brain in a smarter way.

If you use a bad strategy, you may not learn—even if you try hard. A few people study for math by doing the same set of easy problems and skipping the hard ones, or just re-reading the textbook, because it feels easier. Yet when it comes time to do the test, they don't do well because they didn't work on problems that stretched their brains and taught them new things. When this happens, they may even say "I'm just not smart at math."

But the truth is that everyone can become smarter at math if they practice in the right way. If a weight lifter watched other people exercise all day long, he wouldn't get any stronger. And if someone tried to learn how to juggle by just reading a book about juggling, they wouldn't learn. You actually have to practice the right way—and usually that means the hard way—to get better at something. In fact, scientists have found that the brain grows more when you learn something new, and less when you practice things you already know.

This means that it's not just how much time and effort you put in to studying math, but whether, when you study, you learn something new and hard. To do that, you usually need to use the right strategies. People often learn those good strategies from others, like teachers or students who do well. Luckily, strategies are easy to learn if you get help.

The Truth About "Smart" and "Dumb"

People aren't "smart" or "dumb" at math. At first, no one can read or solve equations. But with practice, they can learn to do it. And the more a person learns, the easier it gets to learn new things—because their brain "muscles" have gotten stronger.

This is true even for adults who have struggled for a long time to learn something. Dr. Wittenberg, a scientist from Wake Forest University, said "We used to think adults can't form new brain connections, but now we know that isn't true... The adult brain is like a muscle, and we need to exercise it."

People who don't know this can miss out on the chance to grow a stronger brain. They may think they can't do it, or that it's too hard. It does take work to learn, just like becoming stronger physically or becoming a better juggler does. Sometimes it even hurts! But when you feel yourself get better and stronger, you realize that all the work is worth it!

References:

A similar version of this article was written by Lisa Blackwell and can be downloaded from: www.brainology.us/websitemedia/youcangrowyourintelligence.pdf

Blackwell, L. A., Trzesniewski, K. H., & Dweck, C. S. (2007). Theories of intelligence and achievement across the junior high school transition: A longitudinal study and an intervention. *Child Development, 78*, 246-263.

Driemeyer, J., Boyke, J., Gaser, C., Buchel, C., May, A. (2008). Changes in Gray Matter Induced by Learning—Revisited. *PLoS One, 3*, e2669.
doi:10.1371/journal.pone.0002669.

Nordqvist, C. (2004, Feb 1). “Juggling makes your brain bigger – New Study.” Retrieved from <http://www.medicalnewstoday.com/releases/5615.php>

Appendix 2: Growth Mindset Scale [Adapted from [28]]

Respond to the following questions based on the below scale:

1 = strongly agree

2 = agree

3 = mostly agree

4 = mostly disagree

5 = disagree

6 = strongly disagree.

Questions

You have a certain amount of scientific research skill, and you can't really do much to change it.

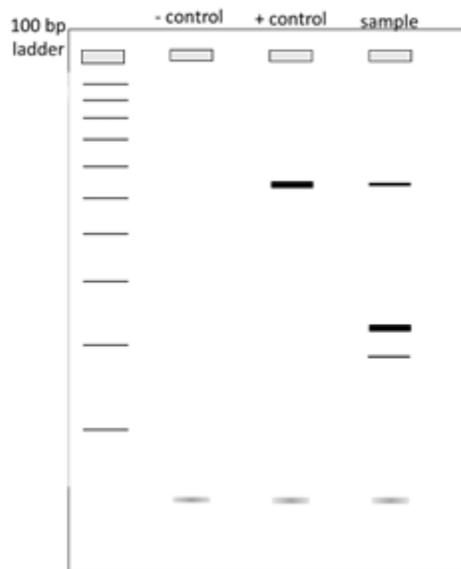
Your science research ability is something about you that you can't change very much.

You can learn new things, but you can't really change your scientific research skills.

Appendix 3: Post intervention Quiz: [Adapted from [37]].

1) What does PCR stand for? (Assessing “Knowledge”)

- a. **Polymerase Chain Reaction**
- b. Polymerase Cellular Rate
- c. Pretty Cool Reaction



To the left is an agarose gel that was run to determine if you successfully amplified a gene of interest (gene X) from bacterial DNA. To amplify gene X you performed a PCR reaction that included your primers, dNTPs, reaction buffer and both a negative and a positive control.

2) What do the bands on this agarose gel represent?
(Assessing “Comprehension”)

- a. Bacterial protein
- b. Primer concentration
- c. Amplified bacterial DNA separated by gene
- d. **Amplified bacterial DNA separate by size**

3) T/F From our gel, we can conclude that gene X is 500-600bp. (Assessing “Analysis”)

4) T/F The band with the highest intensity in our sample lane is amplified gene X. (Assessing “Analysis”)

- 5) In looking at the agarose gel, and thinking through the components of our PCR reaction, what can we conclude about the efficiency of our PCR reaction? Select all that apply (Assessing “Evaluation”)
- a. Our PCR reaction was efficient because we amplified gene X.
 - b. Our PCR reaction is inefficient as we have diffuse small bands in each reaction.**
 - c. Our PCR reaction is inefficient as we have non-specific bands amplified in our sample reaction.**
 - d. Our PCR reaction failed.

Scoring:

Question 1 = 1 point

Question 2 = 2 point

Question 3 = 4 point

Question 4 = 4 points

Question 4 = 18 point (9 points for picking ‘b’, 9 points for picking ‘c’)

Total = 29 points

Appendix 4: Short GRIT-S Scale [34, 35]

Short Grit Scale

Directions for taking the Grit Scale: Please respond to the following 8 items. Be honest – there are no right or wrong answers!

1. New ideas and projects sometimes distract me from previous ones.*
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

2. Setbacks don't discourage me.
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

3. I have been obsessed with a certain idea or project for a short time but later lost interest.*
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

4. I am a hard worker.
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

5. I often set a goal but later choose to pursue a different one.*
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

6. I have difficulty maintaining my focus on projects that take more than a few months to complete.*
 - Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

7. I finish whatever I begin.
- Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

8. I am diligent.
- Very much like me
 - Mostly like me
 - Somewhat like me
 - Not much like me
 - Not like me at all

Scoring:

1. For questions 2, 4, 7 and 8 assign the following points:
 - 5 = Very much like me
 - 4 = Mostly like me
 - 3 = Somewhat like me
 - 2 = Not much like me
 - 1 = Not like me at all

2. For questions 1, 3, 5 and 6 assign the following points:
 - 1 = Very much like me
 - 2 = Mostly like me
 - 3 = Somewhat like me
 - 4 = Not much like me
 - 5 = Not like me at all

Add up all the points and divide by 8. The maximum score on this scale is 5 (extremely gritty), and the lowest score on this scale is 1 (not at all gritty).

Grit Scale citation

Duckworth, A.L., & Quinn, P.D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*, 166-174.
<http://www.sas.upenn.edu/~duckwort/images/Duckworth%20and%20Quinn.pdf>

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 9*, 1087-1101.
<http://www.sas.upenn.edu/~duckwort/images/Grit%20JPSP.pdf>

Appendix 5: Science Motivation Questionnaire II (SMQII) [36]

Science Motivation Questionnaire II

In order to better understand what you think and how you feel about your college science courses, please respond to each of the following statements from the perspective of “When I am in a college science course. . .”

[Response scale: Never Rarely Sometimes Usually Always]

01. The science I learn is relevant to my life
 02. I like to do better than other students on science tests
 03. Learning science is interesting
 04. Getting a good science grade is important to me
 05. I put enough effort into learning science
 06. I use strategies to learn science well
 07. Learning science will help me get a good job
 08. It is important that I get an “A” in science
 09. I am confident I will do well on science tests
 10. Knowing science will give me a career advantage
 11. I spend a lot of time learning science
 12. Learning science makes my life more meaningful
 13. Understanding science will benefit me in my career
 14. I am confident I will do well on science labs and projects
 15. I believe I can master science knowledge and skills
 16. I prepare well for science tests and labs
 17. I am curious about discoveries in science
 18. I believe I can earn a grade of “A” in science
 19. I enjoy learning science
 20. I think about the grade I will get in science
 21. I am sure I can understand science
 22. I study hard to learn science
 23. My career will involve science
 24. Scoring high on science tests and labs matters to me
 25. I will use science problem-solving skills in my career
- End. Thank you
-

Factor 1. Intrinsic motivation

Learning science is interesting
I am curious about discoveries in science
The science I learn is relevant to my life
Learning science makes my life more meaningful
I enjoy learning science

Factor 2. Career motivation

Learning science will help me get a good job
Understanding science will benefit me in my career
Knowing science will give me a career advantage
I will use science problem-solving skills in my career
My career will involve science

Factor 3. Self-determination

I study hard to learn science
I prepare well for science tests and labs
I put enough effort into learning science
I spend a lot of time learning science
I use strategies to learn science well

Factor 4. Self-efficacy

I believe I can earn a grade of "A" in science
I am confident I will do well on science tests
I believe I can master science knowledge and skills
I am sure I can understand science
I am confident I will do well on science labs and projects

Factor 5. Grade motivation

Scoring high on science tests and labs matters to me
It is important that I get an "A" in science
I think about the grade I will get in science
Getting a good science grade is important to me
I like to do better than other students on science tests

Students respond to each item on a rating scale: never (0), rarely (1), sometimes (2), often (3), or always

(4). There is a raw score of 20 points for each factor.

A3.7 References

1. Henry, M.A., et al., *FAIL Is Not a Four-Letter Word: A Theoretical Framework for Exploring Undergraduate Students' Approaches to Academic Challenge and Responses to Failure in STEM Learning Environments*. Cbe-Life Sciences Education, 2019. **18**(1): p. 17.
2. Indorf, J.L., et al., *Adding Authenticity to Inquiry in a First-Year, Research-Based, Biology Laboratory Course*. CBE Life Sci Educ, 2019. **18**(3): p. ar38.
3. Marra, R.M., et al., *Leaving Engineering: A Multi-Year Single Institution Study*. Journal of Engineering Education, 2012. **101**(1): p. 6-27.
4. Simpson, A. and A. Maltese, "Failure Is a Major Component of Learning Anything": *The Role of Failure in the Development of STEM Professionals*. Journal of Science Education and Technology, 2017. **26**(2): p. 223-237.
5. *Integrated postsecondary education data system completions survey*, W. <https://caspar.nsf.gov/>. Editor. 2010.
6. Maltese, A.V. and R.H. Tai, *Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students*. Science Education, 2011. **95**(5): p. 877-907.
7. Laursen, S., Hunter, A.B., Seymour, E., Thiry, H., Melton, G., *Undergraduate research in the sciences: Engaging students in real science*. 2010: John Wiley & Sons.
8. Lopatto, D., et al., *Undergraduate research. Genomics Education Partnership*. Science, 2008. **322**(5902): p. 684-5.
9. Thiry, H., et al., *The benefits of multi-year research experiences: differences in novice and experienced students' reported gains from undergraduate research*. CBE Life Sci Educ, 2012. **11**(3): p. 260-72.
10. Dweck, C.S., *Self-theories: their role in motivation, personality, and development*. Essays in social psychology. 2000, Philadelphia, Pa.: Psychology Press. 195.

11. Dweck, C.S., *Mindset: the new psychology of success*. Ballantine Books trade pbk. ed ed. 2008, New York: Ballantine Books. 277.
12. Ehrlinger, J., A.L. Mitchum, and C.S. Dweck, *Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment*. *Journal of Experimental Social Psychology*, 2016. **63**: p. 94-100.
13. Haimovitz, K. and C.S. Dweck, *The Origins of Children's Growth and Fixed Mindsets: New Research and a New Proposal*. *Child Dev*, 2017. **88**(6): p. 1849-1859.
14. Lou, N.M. and K.A. Noels, *Measuring Language Mindsets and Modeling Their Relations With Goal Orientations and Emotional and Behavioral Responses in Failure Situations*. *The Modern Language Journal*, 2017. **101**(1): p. 214-243.
15. *The Perils and Promises of Praise - Educational Leadership*.
16. Forsythe, A. and S. Johnson, *Thanks, but no-thanks for the feedback*. *Assessment & Evaluation in Higher Education*, 2017. **42**(6): p. 850-859.
17. Clifford, M.M., A. Kim, and B.A. McDonald, *Responses to Failure as Influenced by Task Attribution, Outcome Attribution, and Failure Tolerance*. *The Journal of Experimental Education*, 1988. **57**(1): p. 17-37.
18. Heine, S.J., et al., *Divergent consequences of success and failure in japan and north america: an investigation of self-improving motivations and malleable selves*. *J Pers Soc Psychol*, 2001. **81**(4): p. 599-615.
19. Brdar, I., M. Rijavec, and D. Loncaric, *Goal orientations, coping with school failure and school achievement*. *European Journal of Psychology of Education*, 2006. **21**(1): p. 53-70.
20. Mortenson, S.T. *Cultural Differences and Similarities in Seeking Social Support as a Response to Academic Failure: A Comparison of American and Chinese College Students*. Taylor & Francis [doi:10.1080/03634520600565811]. 2006.
21. Shin, H., et al., *Relationships between coping strategies and burnout symptoms: A meta-analytic approach*. *Professional Psychology: Research and Practice*, 2014. **45**(1): p. 44-56.

22. Smiley, P.A., et al., *Mediation models of implicit theories and achievement goals predict planning and withdrawal after failure*. *Motivation and Emotion*, 2016. **40**(6): p. 878-894.
23. Snyder, K.E., et al., *The message matters: The role of implicit beliefs about giftedness and failure experiences in academic self-handicapping*. *Journal of Educational Psychology*, 2014. **106**(1): p. 230-241.
24. Robins, R.W. and J.L. Pals, *Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change*. *Self and Identity*, 2002. **1**(4): p. 313-336.
25. Dai, T. and J.G. Cromley, *Changes in implicit theories of ability in biology and dropout from STEM majors: A latent growth curve approach*. *Contemporary Educational Psychology*, 2014. **39**(3): p. 233-247.
26. *NSF Award Search: Award#0535903 - Developing Bean Beetles as a Model System for Undergraduate Laboratories*.
27. cbeck, *A Handbook on Bean Beetles, Callosobruchus maculatus – Bean Beetles*.
28. Burnette, J.L., et al., *A Growth Mind-Set Intervention Improves Interest but Not Academic Performance in the Field of Computer Science*. *Social Psychological and Personality Science*, 2019. **11**(1): p. 107-116.
29. Yeager, D.S., et al., *Using Design Thinking to Improve Psychological Interventions: The Case of the Growth Mindset During the Transition to High School*. *J Educ Psychol*, 2016. **108**(3): p. 374-391.
30. Blackwell, L.S., K.H. Trzesniewski, and C.S. Dweck, *Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention*. *Child Development*, 2007. **78**(1): p. 246-263.
31. Crano, W.D. and R. Prislin, *Attitudes and Persuasion*. *Annual Review of Psychology*, 2005. **57**(1): p. 345-374.
32. Morgenroth, T., M.K. Ryan, and K. Peters, *The Motivational Theory of Role Modeling: How Role Models Influence Role Aspirants' Goals*. *Review of General Psychology*, 2015. **19**(4): p. 465-483.

33. Burnette, J.L. and E.J. Finkel, *Buffering against weight gain following dieting setbacks: An implicit theory intervention*. *Journal of Experimental Social Psychology*, 2012. **48**(3): p. 721-725.
34. Duckworth, A.L. and P.D. Quinn, *Development and Validation of the Short Grit Scale (Grit-S)*. *Journal of Personality Assessment*, 2009. **91**(2): p. 166-174.
35. Duckworth, A.L., et al., *Grit: Perseverance and passion for long-term goals*. *Journal of Personality and Social Psychology*, 2007. **92**(6): p. 1087-1101.
36. Glynn, S.M., et al., *Science motivation questionnaire II: Validation with science majors and nonscience majors*. *Journal of Research in Science Teaching*, 2011. **48**(10): p. 1159-1176.
37. Crowe, A., C. Dirks, and M.P. Wenderoth, *Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology*. *CBE—Life Sciences Education*, 2008. **7**(4): p. 368-381.