Nuclear accumulation of polyadenylated noncoding RNA leads to a breakdown in nuclear RNA homeostasis

by

Biplab Paul

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Cell Biology

University of Alberta

Abstract

The biogenesis and processing of messenger RNAs (mRNAs) is central to the gene expression program. In eukaryotic cells, mRNAs must also be exported from nucleus to cytoplasm as part of the gene expression pathway. In this work, mutations in 1047 essential genes of *Saccharomyces cerevisiae* (*S. cerevisiae*) were screened for defects in mRNA export resulting in the identification of 26 mRNA export mutants. Single molecule fluorescence *in situ* hybridization (smFISH) experiments further showed that these mutants accumulated mRNAs within specific regions of the nucleus, which included: (1) transcripts enriched near nuclear pore complexes when components of the mRNA export apparatus were mutated, (2) build-up of mRNAs near transcription sites associated with mutations in genes required for 3′ end processing and chromosome segregation, and (3) mRNAs within the nucleolus when nucleocytoplasmic transport (e.g. *srm1-ts*), rRNA biogenesis (e.g. *enp1-1*), or RNA processing and surveillance (e.g. *csl4-ph*) were disrupted. These data demonstrate that alterations to RNA processing and overall nuclear function cause RNAs to stall, or be retained, at three distinct restriction points. This may reflect common failures in mRNA biogenesis and export, as well as active mechanisms to hold mRNAs at discrete locations to protect the cell and fidelity of the gene expression when cells are dysfunctional.

The nucleolus is mainly associated with noncoding RNA (ncRNA) processing, raising questions about the exact nature of the accumulated poly(A)-RNA material in *csl4-ph*, *enp1-1*, and *srm1-ts* strains. RNA-sequencing (RNA-seq) methods showed that these three mutants exhibited significant down-regulation of protein-coding transcripts that are highly expressed under normal steady state growth conditions, but they also exhibited increased levels of pervasive transcripts. Combined RNA-seq, Northern, and RNA binding data, further revealed that the errors

in RNA biogenesis in *csl4-ph* and *enp1-1* mutants led to the accumulation of polyadenylated ncRNA species. Loss of Csl4p or Enp1p was also found to result in a poly(A)-RNA binding protein, Nab2p, engaging in a protein interaction network that encompassed ncRNA processing factors. This included Nab2p becoming associated with proteins required for ribosomal RNA (rRNA) and small nucleolar RNA (snoRNA) processing, in contrast to the normal association of Nab2p with mRNA processing factors in a control strain. These data indicated that ncRNAs become stably polyadenylated in *csl4-ph* and *enp1-1* mutants, which leads to an inappropriate association between other mRNA processing factors and ncRNAs. In line with this model, overexpression of another poly(A)-RNA binding protein, Pab1p, rescued the poly(A)-RNA accumulation phenotype and improved the growth of the *enp1-1* strain at a semi-permissive temperature. These findings provide evidence that polyadenylated ncRNAs can sequester mRNA biogenesis and export machinery, which precipitates a breakdown in nuclear homeostasis defined by high levels of nuclear poly(A)-RNA in the nucleolus. A set of events that can be mitigated by the overexpression of another poly(A)-RNA binding protein.

Overall, these studies have identified essential genes that are required for the maintenance of the gene expression program and provided new insight into nuclear RNA homeostasis. A key finding is that mutations altering processes not directly linked to mRNA biogenesis or export (i.e. Enp1p) have the ability to disrupt gene expression through altering the abundance of functional RNA-binding proteins. Data that highlights the importance of nuclear surveillance and decay of aberrant RNA species in maintaining nuclear RNA homeostasis and proper cellular function.

Preface

Chapter III of this thesis has been previously published as: **Paul, B., and Montpetit, B. (2016). Altered RNA processing and export lead to retention of mRNAs near transcription sites and nuclear pore complexes or within the nucleolus. Mol. Biol. Cell *27*, 2742–2756.** I performed all the described experiments related to screening the mutant collection and characterizing mRNA export mutants. I was also responsible for data analyses, figure generation, and manuscript editing. Dr. Ben Montpetit was the supervisory author and was involved with original concept formation, manuscript composition, manuscript editing, and data analyses.

Portions of Chapter IV of this thesis are currently in preparation to be submitted for publication as **Paul, B., Aguilar, LC., Gendron, L., Rajan, AAR., Montpetit, R., Trahan, C., Pechmann, S., Oeffinger, M. and Montpetit. B. Stabilization of polyadenylated noncoding RNA species leads to an overall disruption in nuclear RNA homeostasis.** I was responsible for data collection and analyses, and original manuscript writing. Carolina Aguilar (Institut de Recherches Cliniques de Montréal) and I contributed equally to this work as co-first authors. Carolina Aguilar assisted with data collection and analyses, which I have indicated in the corresponding figures. Drs. Ben Montpetit, Marlene Oeffinger (Institut de Recherches Cliniques de Montréal), and Sebastian Pechemann (Université de Montréal) were the supervisory authors and were also involved with concept formation, manuscript composition, and data analyses.

Dedication

To my parents for their love and support…

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Dr. Ben Montpetit, for sharing his unwavering passion for science. It has been an honor to be one of his first Ph.D. students. I appreciate all of his contributions of time, ideas, and funding to make my Ph.D. productive and stimulating. I am thankful for providing me appropriate training and advice whenever needed. I also would like to thank him for showing me the path of high standard and excellence which I will continue to strive for. I am also thankful to my committee members Drs. Richard Rachubinski and Martin Srayko for their continuous support and guidance throughout my graduate school.

I would like to thank a really fabulous lab manager Rachel Montpetit for her support throughout my journey at UC Davis, which made my graduate school experience as smooth as possible. I am also grateful for the friendship of Azra Lari and Kelly Tedrick for sharing their knowledge of yeast genetics with me, which was tremendously helpful throughout my Ph.D. program. I would like to thank other past and present members of the Montpetit laboratory that I have had pleasure to work with, or alongside of, including Matt McDermand, Arvind Arul Nambi Rajan, Judit Voros, Rima Sandhu, Ryuta Asada, and summer students Benedict Yong, Caitlin Porter, and Iris Unterweger.

Lastly, I would like to thank my friends and family for all their love and encouragement. For my parents who showed me a love for science and supported all my endeavors. Most of all for my loving, supportive, encouraging and patient wife Apu Roy, whose love and support throughout my graduate school is so much appreciated.

Table of Contents

List of Tables

List of Figures

List of Symbols, Abbreviations, and Nomenclature

%	Percent
°C	Degree Celsius
AP-MS	Affinity Purification Mass Spectrometry
APA	Alternative Polyadenylation
ATP	Adenosine Triphosphate
bp	base pair
CBC	Cap Binding Complex
CF	Cleavage Factor
ChIP	Chromatin Immunoprecipitation
CPF	Cleavage and Polyadenylation Factor
CTD	Carboxy Terminal Domain
CUT	Cryptic Unstable Transcript
Da	Dalton
DAmP	Decreased Abundance by mRNA Perturbation
DAPI	4’ 6 diamidino-2-phenylidole
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic Acid
EJC	Exon Junction Complex
ePAT	extended Polyadenylation Test
ETS	External Transcribed Spacers
FC	Fold Change
FG	Phenylalanine-Glycine
FISH	Fluorescence In Situ Hybridization
FITC	Fluorescein IsoThioCyanate
g	gravitational acceleration
G1	Gap1
GFP	Green Fluorescent Protein
GO	Gene Ontology
HDAC	Rpd3S Histone Deacetylase
IgG	Immunoglobulin G
IP	Immunoprecipitation
ITS	Internal Transcribed Spacer
ITS1	Internal Transcribed Spacer 1
ITS2	Internal Transcribed Spacer 2
kbp	kilo base pair
KH	K-Homology
L	Liter
LNA	Locked Nucleic Acid
m	Meter
M	Molarity
M-phase	Mitotic-phase
mg	milligram
min	minutes
ml	milliliter
mM	millimolar

mRNA	messenger Ribonucleic Acid
mRNP	messenger Ribonucleoparticle
ms	millisecond
NA	Numerical Aperture
ncRNA	noncoding Ribonucleic Acid
NDR	Nucleosome Depleted Region
NET-Seq	Native Elongating Transcript Sequencing
ng	Nanogram
nL	Nanoliter
nm	Nanometer
NMD	Nonsense Mediated Decay
NNS	Nrd1-Nab3-Sen1
NPC	Nuclear Pore Complex
NTP	Nucleoside Triphosphate
Nup	Nucleoporin
NUT	NRD1 Unstable Transcripts
OD	Optical Density
ORF	Open Reading Frame
PABP	Poly(A) Binding Protein
PAGE	Polyacralymide Gel Electrophoresis
PEG	Polyethelene Glycol
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PH	Phosphorolysis domain
poly(A)	Polyadenyl
ppm	parts per million
RBP	RNA binding protein
RCF	Relative Centrifugal Force
RNA	Ribonucleic Acid
RNA-seq	RNA sequencing
RNAP I	RNA Polymerase I
RNAP II	RNA Polymerase II
RNAP III	RNA Polymerase III
RNP	Ribonucleoparticle
RP	Ribosomal Protein
RPKM	Read per Kilobase Per Million
RRM	RNA Recognition Motif
rRNA	ribosomal RNA
RT	Room Temperature
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SC	Synthetic Complete
SDS	Sodium Dodecyl Sulphate
SKI	Super Killer
smFISH	single molecule FISH
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA

SR	Serine Arginine
R	R programming language
SSC	Saline-Sodium Citrate
ssDNA	Salmon sperm DNA
SUT	Stable Unannotated Transcripts
TES	Transcription End Site
TRAMP	Trf4/Air2/Mtr4 Polyadenylation
TREX	Transcription Export
tRNA	Transfer RNA
Ts	Temperature sensitive
TSC	Total Spectral Counts
TSS	Transcription Start Site
µl	microliter
UTR	Untranslated Region
UV	Ultra-Violet
VRC	Vanadyle Ribonucleoside Complex
WT	Wild Type
XUT	XRN1 Unstable Transcripts
YPD	Yeast Extract–Peptone–Dextrose

Chapter I: *Introduction*

1.1 General eukaryotic RNA biogenesis and metabolism

The function and survival of an individual cell is dependent on the ability of the cell to appropriately employ information stored in DNA. The specific information being used is different for each cell and can change in response to physiological and environmental stimuli. This is mainly accomplished through changes in the gene expression program, which drives changes at the level of both cell structure and function. In eukaryotic cells, the gene expression program encompasses the flow of genetic information from DNA that is stored in the nucleus, to an RNA intermediate, which in most instances the information encoded within the RNA is then used to produce a protein within the cytoplasm (i.e. the central dogma of biology). Given this process, the biogenesis of RNA intermediates must occur in an accurate and efficient manner to allow the flow of genetic information, changes in gene expression, and overall cellular function.

In eukaryotic cells, transcription of RNAs from a DNA template occurs through the activity of three RNA polymerases (RNAPs) that are distinguished by their subunit composition and the transcript they produce (Roeder and Rutter, 1969, 1970). The classes of RNAs synthesized in the cell not only include protein coding, or messenger RNAs (mRNAs), but also noncoding RNAs (ncRNAs), which provide both structural and functional elements required for protein synthesis and other cellular processes (Nissen *et al.*, 2000; Spahn *et al.*, 2001; Selmer *et al.*, 2006; Passmore *et al.*, 2007; Armache *et al.*, 2010; Ben-Shem *et al.*, 2010). ncRNAs can be classified into several different groups, ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNA (snoRNAs), with each of these classes having a defined function within the gene expression program (Balakin *et al.*, 1996; Tycowski *et al.*, 1996; Kiss-László *et al.*, 1998; Yusupov *et al.*, 2001; Schmeing *et al.*, 2003; Selmer *et al.*, 2006; Fica *et al.*, 2013; Jorjani *et al.*, 2016). mRNAs, snoRNAs, and some snRNAs are synthesized by RNA Polymerase II (RNAPII), rRNAs are synthesized by both RNA polymerase I (RNAPI) and RNA polymerase III (RNAPIII), and tRNAs and snRNAs are synthesized by RNA polymerase III (RNAPIII) (Weinmann and Roeder, 1974; Dignam *et al.*, 1983). A noteworthy consequence of the transcriptional process is the generation of pervasive transcripts, which are thought to largely arise from spurious RNAPII transcription and in most cases generate transcripts with no cellular function (Velculescu *et al.*, 1997; Kaplan *et al.*, 2003; Wyers *et al.*, 2005; Kapranov *et al.*, 2007; Neil *et al.*, 2009; Xu *et al.*, 2009).

Newly transcribed RNA transcripts undergo various processing and modification steps, which occur co-transcriptionally in some instances, but also occur post-transcriptionally in both the nuclear and cytoplasmic compartments (Figure 1-1). RNA biogenesis and processing are mediated by RNA-binding proteins (RBPs) associated with the transcript in the form of a ribonucleoprotein particle (RNP), with individual RBPs defining the path taken by the RNP through RNA metabolism (Amrani *et al.*, 1997; Kessler *et al.*, 1997; Zhao *et al.*, 1999; Hector *et al.*, 2002; Lemay *et al.*, 2010). As such, the composition of the RNP in terms of the RBPs that are present varies throughout the RNA life cycle, beginning from transcription until it is turned over through decay processes in the nuclear or cytoplasmic compartments (Mitchell *et al.*, 2013). Given the complexity of RNP biogenesis, the generation of aberrant RNAs is inevitable due to the potential for an error at any stage of RNA metabolism. Consequently, surveillance for and decay of aberrant RNAs are critical processes that act to prevent entry of non-functional RNA substrates into the gene expression program and the accumulation of stalled processing intermediates (Figure 1-1). Proper functioning of the surveillance and decay machinery is paramount to the ability of a cell to maintain nuclear RNA homeostasis, which is defined here as the proper and ongoing biogenesis, processing, and production of functional RNA species that are required to support gene expression and cellular function.

1.2 Nuclear mRNA processing

A large component of RNA metabolism includes the biogenesis, processing, and export of messenger-RNPs or mRNPs. There are several nuclear processing stages required to generate export competent mRNPs, and importantly these processes are often linked such that they occur in a coupled manner. Linking processing steps is likely important for the timely and efficient generation of export competent mRNPs, as well as a mechanism to ensure proper mRNP assembly (Oeffinger and Montpetit, 2015). Given that the majority of spurious transcription leading to the generation of pervasive transcripts occurs through RNAPII activity, linking these processes may further ensure that pervasive transcripts do not continue in the gene expression pathway. However, as of now it is not fully understood how mRNA biogenesis is differentiated from pervasive transcription since they both are generated by RNAPII.

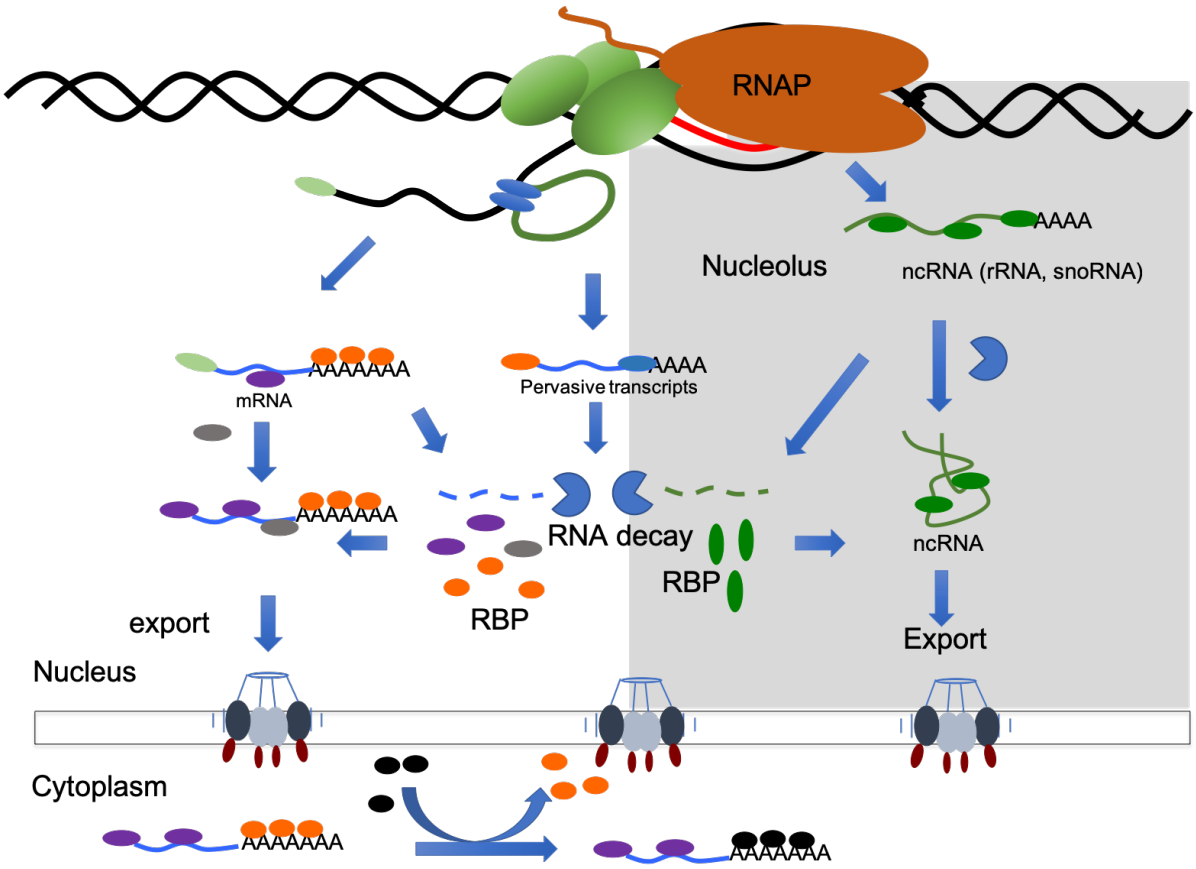


Figure 1-1. RNA processing in the nucleus. Both coding (mRNA) and ncRNA (rRNA, snoRNA, pervasive transcripts) are synthesized in the nucleus. In the nucleoplasm, associated with transcription, mRNAs undergo various processing steps such as capping, splicing, and 3′-end processing to be followed by export to the cytoplasm. In contrast, most pervasive transcripts are rapidly degraded following transcription within the nucleus by RNA decay machineries. For ncRNAs, such as rRNAs and snoRNAs, transcripts are synthesized and/or processed in association with the nucleolus (gray area). By-products produced from rRNA processing are also degraded by RNA decay machineries. After processing, ncRNAs are exported to the cytoplasm to participate in protein synthesis and other cellular processes. When an error in processing occurs, an RNA can be subject to degradation by the nuclear RNA decay machinery. As a result of RNA decay, bound RBPs are released from these aberrant RNAs to be recycled back to the various RNA processing pathways. Given the sharing of common RBPs, the processing of all classes of RNA are interconnected, with defects in any one pathway having the potential to feedback and disrupt other RNA biogenesis pathways.

In the case of mRNA transcription, mRNP formation begins co-transcriptionally, mediated by RNAPII itself (Moteki and Price, 2002). This includes co-transcriptional processing events coordinated by the carboxy terminal domain (CTD) of the large subunit of RNAPII (Allison *et al.*, 1988; Hsin and Manley, 2012). For example, it is known that reversible phosphorylation of the CTD couples different steps of transcription, including initiation, elongation, and termination, to RNA processing via recruitment of RBPs to the transcribing RNA (Lu *et al.*, 1991; Conaway *et al.*, 1992). Nuclear mRNA processing steps further include 5′ capping, splicing, 3′ end processing, polyadenylation, and assembly of a mature mRNP that can be exported to the cytoplasm via nuclear pore complexes (NPCs) (Figure 1-2). Once in the cytoplasm, mRNAs can be used for protein synthesis, transported to specific locations in the cytoplasm, stored for later use, or decayed (Thermann *et al.*, 1998).

A near universal nuclear modification to mRNAs that is coupled to transcription is the addition of a 5′ 7-methylguanosine cap. Upon transcription initiation and incorporation of the first 20-30 nucleotides, a 7-methylguanosine cap is added to the 5′ end of the pre-mRNA to protect the transcript from degradation (Coppola *et al.*, 1983). The three enzymes required to generate the cap structure on the mRNA interact with the CTD of RNAPII to couple transcription with 5′ capping (Figure 1-2) (Itoh *et al.*, 1987; Yue *et al.*, 1997; Cho *et al.*, 1998). Cet1p removes the gamma phosphate from the 5′ triphosphate end of the RNA, a guanosine monophosphate (GMP) is then added to the diphosphate end of the RNA by Ceg1p with a 5′-5′ phosphate linkage (Shibagaki *et al.*, 1992; Tsukamoto *et al.*, 1997; Rodriguez *et al.*, 1999). Finally, the 7N methyl group is added to the guanine base by the guanine 7-methyltransferase, Abd1p (Mao *et al.*, 1995). In *S. cerevisiae* there are also two proteins, Rai1p and Rat1p, that possess pyrophosphatase and decapping activity that are responsible for quality control of 5′ capped mRNA (Jiao *et al.*, 2010). Upon capping, the pre-mRNA cap is bound by the cap binding complex (CBC) that consists of Cbp20p and Cbp80p, which can act to initiate splicing through engagement of snRNPs (Black and Steitz, 1986; Izaurralde *et al.*, 1995; Fortes *et al.*, 1999). In line with this, the depletion of CBC proteins is known to alter splicing of pre-mRNAs (Izaurralde *et al.*, 1994). The CBC has also been shown to interact with various export factors to facilitate mRNA export to the cytoplasm (Nojima *et al.*, 2007). Following export, the 5′ cap facilitates translation through association with translation initiation factors, including eIF4G and the DEAD-box RNA helicase eIF4A (Gingras *et al.*, 1999). Through this process, the CBC is replaced by eIF4E, which

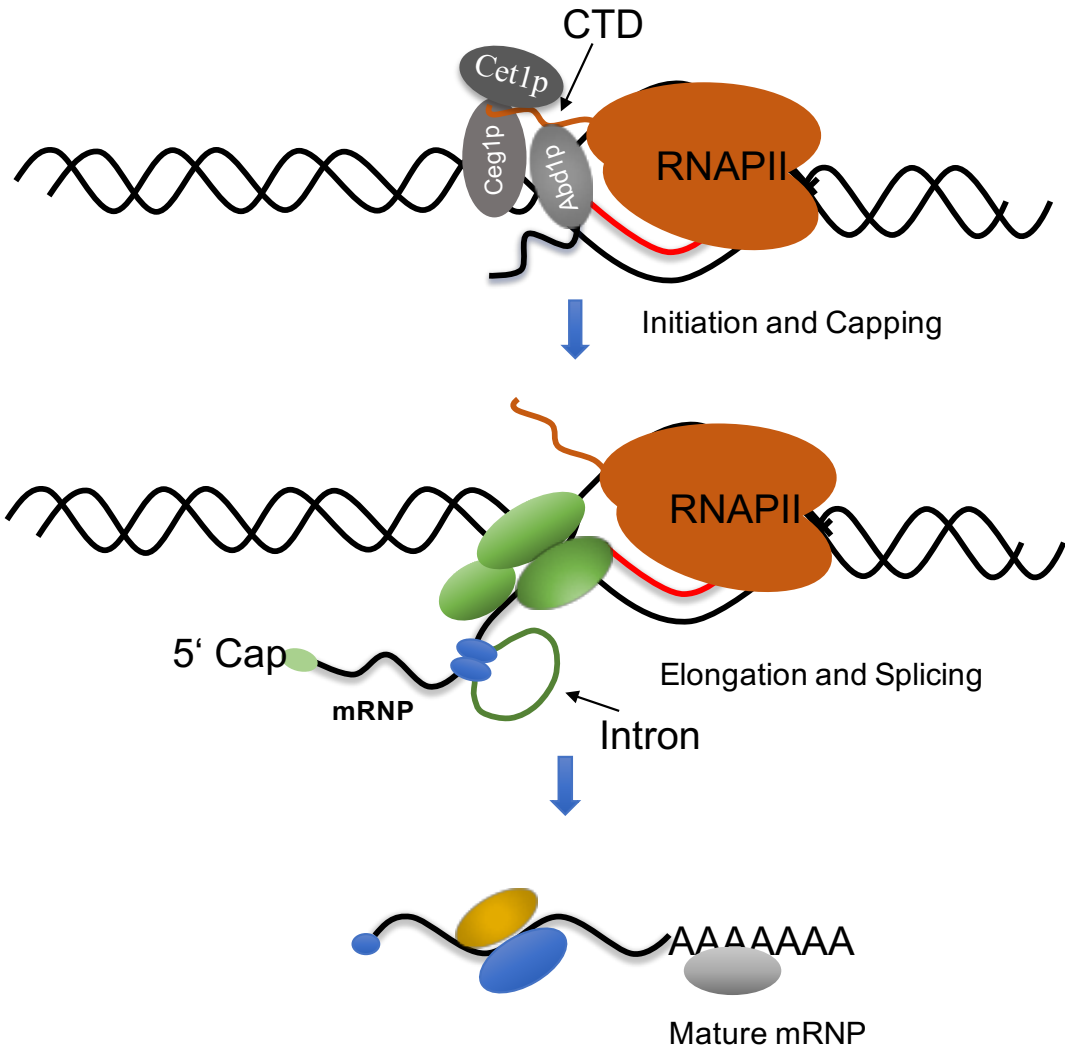


Figure 1-2. Co-transcriptional mRNA processing. Simplified schematic showing the overall nuclear processing of mRNAs. After synthesis of 20-30 nucleotides of the pre-mRNA, enzymes such as Cet1p, Ceg1p, and Abd1p interact with the carboxy terminal domain(CTD) of RNAPII to perform the chemical reactions required to add a 7-methylguanosine cap to the 5′ end of nascent mRNA. During the elongation step of transcription, introns are removed through interactions with splicing machinery, followed by cleavage and polyadenylation to release the mature mRNP for export, which is directed through the actions of various RNA-binding proteins (indicated by green, yellow, and blue circles).

interacts with eIF4G to promote translation (Fortes *et al.*, 2000). Notably, these multiple functions of the CBC exemplify the close connection that RBPs provide between nuclear mRNA biogenesis and downstream processes, which together act to couple transcription, export, and translation.

Many eukaryotic mRNAs contain one or more noncoding intervening sequences, or introns, which need to be removed from pre-mRNA in order to synthesize a functional protein. In *S. cerevisiae*, only 3% of genes contain an intron but these genes account for ~25% of cellular mRNAs, thus splicing is a critical nuclear mRNA processing step for a large pool of mRNAs in budding yeast (Ares *et al.*, 1999; Lopez and Séraphin, 1999). The process of removing an intron from a pre-mRNA, referred to as splicing, occurs co-transcriptionally in the nucleus. Introns are removed through two transesterification reactions using conserved sequences at the 5′ and 3′ ends of the intron referred to as splice sites (Konarska *et al.*, 1985). The process of splicing starts with recognition of the 5′ splice site, 3′ splice site, and branchpoint by U1, U2AF, and U2 snRNPs, which leads to formation of the pre-spliceosome complex (Black *et al.*, 1985). Subsequently, the interaction of U4/U5 and U6 with the pre-spliceosome reconfigures the RNP complex to position the pre-mRNA so first transesterification reaction can occur (Konarska and Sharp, 1987; Wassarman and Steitz, 1992) (Figure 1-3). After splicing, the exon-exon junction becomes associated with the exon junction complex (EJC) (Le Hir *et al.*, 2000; Bono and Gehring, 2011), which promotes downstream translation through interaction with translation initiation factors (Chazal *et al.*, 2013) or non-sense mediated decay of the mRNA based on presence of the EJC downstream of the mRNA stop codon (Gehring *et al.*, 2003; Chamieh *et al.*, 2008; Ma *et al.*, 2008). The 5′ most exon in a spliced mRNA is known to associate with RBPs, including the TRanscription and Export (TREX) complex, to enhance further splicing events and mRNA export (Le Hir *et al.*, 2001).

Pre-mRNAs also undergo 3′ cleavage and polyadenylation in order to terminate transcription and produce mature mRNAs. Cleavage and polyadenylation occurs ~10-30 nucleotides downstream of a conserved adenosine-uracil (AU)-rich sequence called the cleavage and polyadenylation signal (Proudfoot, 2004), where cleavage and polyadenylation factors (CPF) bind (Chen and Moore, 1992). The core CPF complex is comprised of two different sub-complexes, CFI and CFII, and includes other additional factors involved in 3′ end processing (Chen and Moore, 1992). Upon binding, the CPF recognizes and cleaves a U/GU rich region of

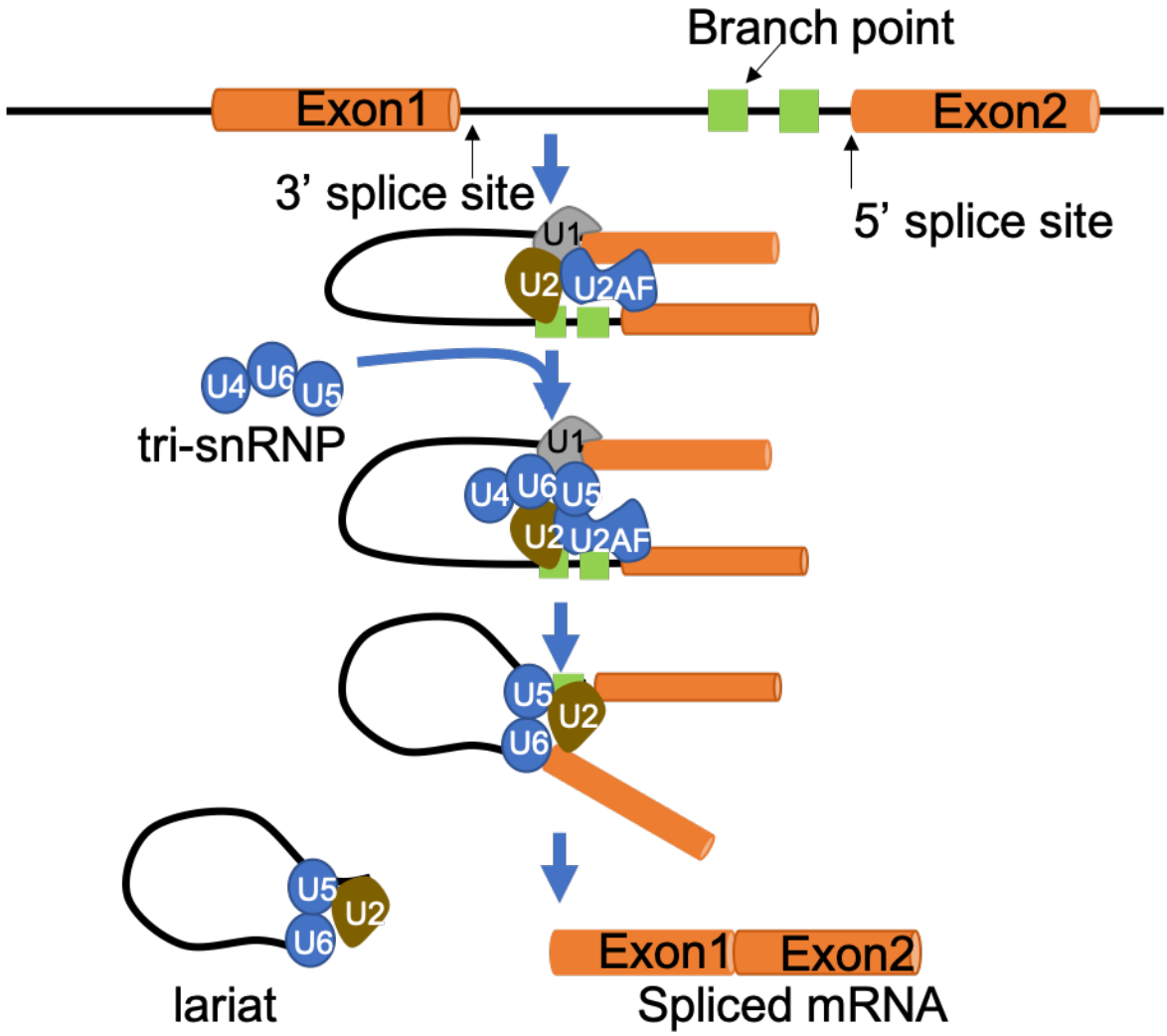


Figure 1-3. Schematic of pre-mRNA splicing. The pre-mRNA intron is shown as a black solid line and exons are represented by cylinders (orange). The U1 and U2 snRNPs recognize the 5′ splice sites and branch point and facilitate the association of the tri-snRNPs (U4, U6, and U5) to this complex, which is then followed by cleavage to remove the intron. This results in the production of an intron lariat and a spliced mRNA. Note only some parts of the spliceosome and mRNA are shown in the figure for simplicity.

the pre-mRNA, which in turn facilitates recruitment of the poly-adenosine (A) polymerase (Pap1p) to add the poly(A) tail to the transcript (Figure 1-4) (Lingner *et al.*, 1991; Kyburz *et al.*, 2003; Ryan *et al.*, 2004). The polyadenylated 3′ ends of mRNAs are reported to be ~70-100 nucleotides long in mammals and 20-40 nucleotides in *S. cerevisiae* (Subtelny *et al.*, 2014; Lima *et al.*, 2017). The recruitment of the protein 1 cleavage polyadenylation factor (Pcf11p) and Rna15p of the CFIA complex to the CTD of RNAPII also facilitates binding of a 5′ → 3′ exonuclease, Xrn2p, to the nascent mRNA to promote transcription termination and release of RNAPII (Sadowski *et al.*, 2003; Kim *et al.*, 2004; Luo *et al.*, 2006; El Hage *et al.*, 2008). In conjunction with cleavage and polyadenylation, in order to protect pre-mRNA from the 3′ to 5′ exonuclease degradation, poly(A) tails become associated with poly(A) binding proteins (PABPs), which also act to promote mRNP export (Fasken *et al.*, 2008; Schmid *et al.*, 2012). Note that mRNAs can be alternatively polyadenylated (APA) in order to synthesize diverse proteins from a fixed number of mRNAs if the APA site is located upstream of the last exon (Hoque *et al.*, 2013). Moreover, alternative 3′ end processing results in the generation mRNAs with different 3′ untranslated regions (UTRs) responsible for altering stability, translation, and localization of transcripts (An *et al.*, 2008; Sandberg *et al.*, 2008; Spies *et al.*, 2013). The release of the mRNP from the site of transcription via 3′ cleavage and polyadenylation allow the mRNP to continue on the gene expression pathway, which involves export to the cytoplasm upon gaining export competence (see below). As indicated for other nuclear processing events discussed above, the recruitment of certain RBPs (e.g. TREX complex and PABPs) during biogenesis is part of the process by which mRNPs become export competent (Strasser *et al.*, 2002; Gwizdek *et al.*, 2005).

1.3 Formation of export competent mRNPs

The presence of a nuclear envelope in eukaryotic cells necessitates that mRNPs must be exported to the cytoplasm to take part in downstream processes such as translation. The overall mechanism and factors required for mRNP export are known to be conserved between yeast and mammals (Strasser *et al.*, 2002). Most mRNP export factors in *S. cerevisiae* have been identified by screening for the accumulation of poly(A)-RNA in the nucleus of temperature sensitive (*Ts*)

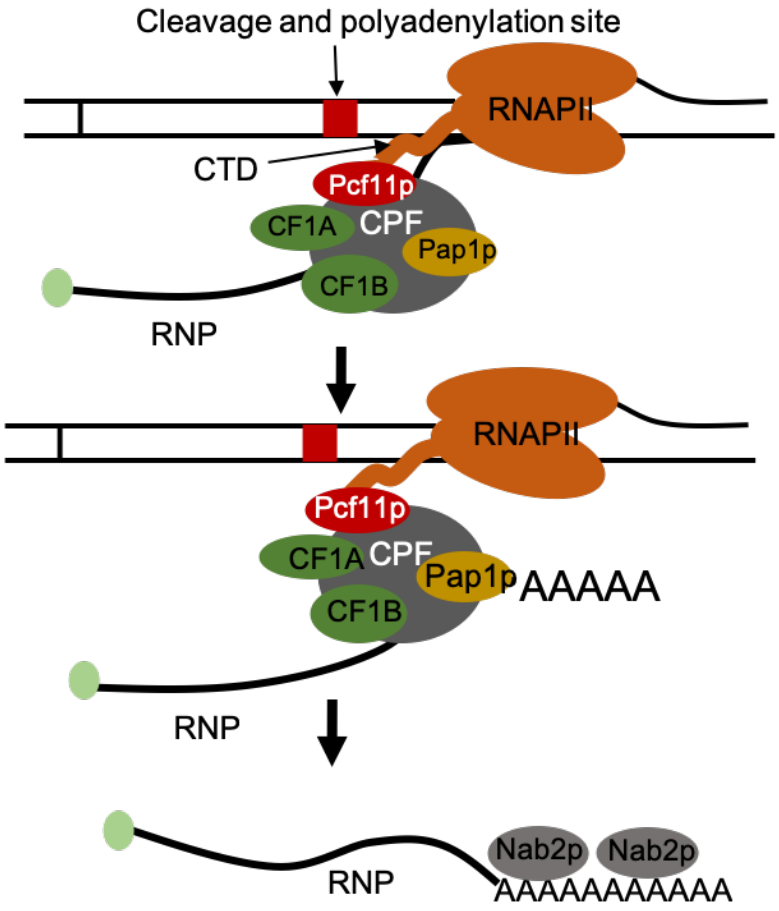


Figure 1-4. Schematic of 3′-end processing. Near the 3′ end of a transcript, cleavage and polyadenylation factors (CPF) and cleavage factors (CF) recognize a polyadenylation signal. After endo-nucleolytic cleavage at the polyadenylation site, a poly(A) tail is added to the 3′ end of the mRNA by poly(A)-polymerase Pap1p (yellow). The nascent polyadenylated mRNA is also bound by Nab2p for protection against nuclear decay. A component of the CPF complex, Pcf11p, interacts with the CTD of RNAPII to facilitate transcription termination.

mutants of essential genes at a non-permissive temperature using fluorescence *in situ* hybridization (FISH) experiments (Amberg *et al.*, 1992; Cole *et al.*, 2002). The identification of these export factors, combined with further characterization of their function, has led to a general model of mRNP assembly leading to export. Briefly, during transcription elongation and 3′ end processing, various proteins factors that function in transcription and export are recruited to the pre-mRNA, leading to the formation of the TREX complex via two steps. In the first step, components of the THO complex, consisting of Tho2p, Hrp1p, Mft1p, and Thp2p, are recruited to the elongating transcript through interactions with both the pre-mRNA and RNAPII (Gewartowski *et al.*, 2012; Peña *et al.*, 2012). In a second step, two protein factors, Yra1p and Sub2p, are transferred to the pre-mRNA (Strabetaer *et al.*, 2000; Strässer *et al.*, 2001; Johnson *et al.*, 2009), which form a larger six protein assembly referred to as the TREX complex (Grzechnik and Kufel, 2008; Rondón *et al.*, 2010). Chromatin immunoprecipitation (ChIP) followed by sequencing (ChIP-Seq) analysis has showed that the TREX complex binds the 3′-end of the RNA, suggesting that the TREX complex is positioned to link transcription elongation, 3′-end processing, and mRNA export (Meinel *et al.*, 2013). An idea that is also supported by protein interactions between the TREX complex and 3′ end processing components (Mandel *et al.*, 2008). In line with these functions of the TREX complex, mutants of the THO/TREX complex exhibit nuclear poly(A)-RNA accumulation and defects in transcription elongation and termination (Libri *et al.*, 2002; Strasser *et al.*, 2002; Rougemaille *et al.*, 2007).

Following TREX assembly, and with polyadenylation, the PABP Nab2p is recruited to form a complex with Yra1p, which subsequently leads to the export factor Mex67p (NXF1 in humans) binding to the mRNP (Katahira *et al.*, 1999; Strabetaer *et al.*, 2000). Of those factors identified, the Mex67p-Mtr2p heterodimer complex acts as the main mRNA export factor in yeast, which is also conserved in other eukaryotes (Segref *et al.*, 1997; Katahira *et al.*, 1999). Upon joining of Mex67p, the mRNP is thought to be export competent as it has been demonstrated that Mex67p interacts with both the mRNA cargo and phenylalanine-glycine nucleoporins (FG-Nups) in order to facilitate mRNP transport through NPCs (Sträßer *et al.*, 2000). Yra1p has also been shown to dissociate from the mRNP just after Mex67 binding, and prior to export, which is thought to be part of a quality control process indicative of the final step of mRNP maturation and a mRNP gaining export competence (Iglesias *et al.*, 2010).

1.4 Current model of mRNP export

Following dissociation of Yra1p, mRNP transport through NPCs is directed by the mRNA export factor Mex67p, which is thought to be one of the last proteins loaded onto the maturing mRNP (Segref *et al.*, 1997; Köhler *et al.*, 2007). This is likely important to prevent immature mRNPs from exiting the nucleus, since Mex67p promotes interactions between the mRNP and components of the nuclear basket of NPCs, such as Nup1p and Nup2p, to facilitate export (Terry and Wente, 2007). At this point, interactions between the PABP Nab2p and the nuclear basket component Mlp1p also promote export, which may further act as a quality control checkpoint of the mRNP assembly state (Hector *et al.*, 2002). In some instances, mRNAs may be transcribed near NPCs as a means to regulate gene expression, which involves tethering a gene to the nuclear periphery in a process referred to as gene-gating (Blobel, 1985). This may be important in the context of highly transcribed or regulated genes to promote their expression and export. Genome-wide ChIP assays having indicated that highly transcribed genes are physically associated with NPC nuclear basket components (Casolari *et al.*, 2004).

After an mRNP docks at the nuclear side of the NPC, transport is facilitated through the interaction of Mex67p with FG-Nups that line the central NPC transport channel (Adams *et al.*, 2014). Following transport through the NPC, the mRNP reaches the cytoplasmic side where NPC components and their binding partners enforce directionality (Strässer *et al.*, 2001; Terry and Wente, 2007). Specifically, the highly conserved DEAD-box protein 5 (Dbp5p) is proposed to displace both Mex67p and Nab2p from the mRNP (Lund and Guthrie, 2005; Tran *et al.*, 2007). Dbp5p activity is stimulated on the cytoplasmic side of NPCs by Nup159p and Gle1p with the small molecule inositol hexakisphosphate (Weirich *et al.*, 2004, 2006; Alcazar-Roman *et al.*, 2006; Montpetit *et al.*, 2011) (Figure 1-5). From the initial docking event to release into the cytoplasm, the time required for export has been measured to be ~200 milliseconds in budding yeast and mammalian cells (Grünwald and Singer, 2010; Smith *et al.*, 2015). Upon mRNP release into the cytoplasm, export factors removed from the mRNP can then be re-imported into the nucleus to participate in another round of mRNA export (Anderson *et al.*, 1993; Lee *et al.*, 1996).

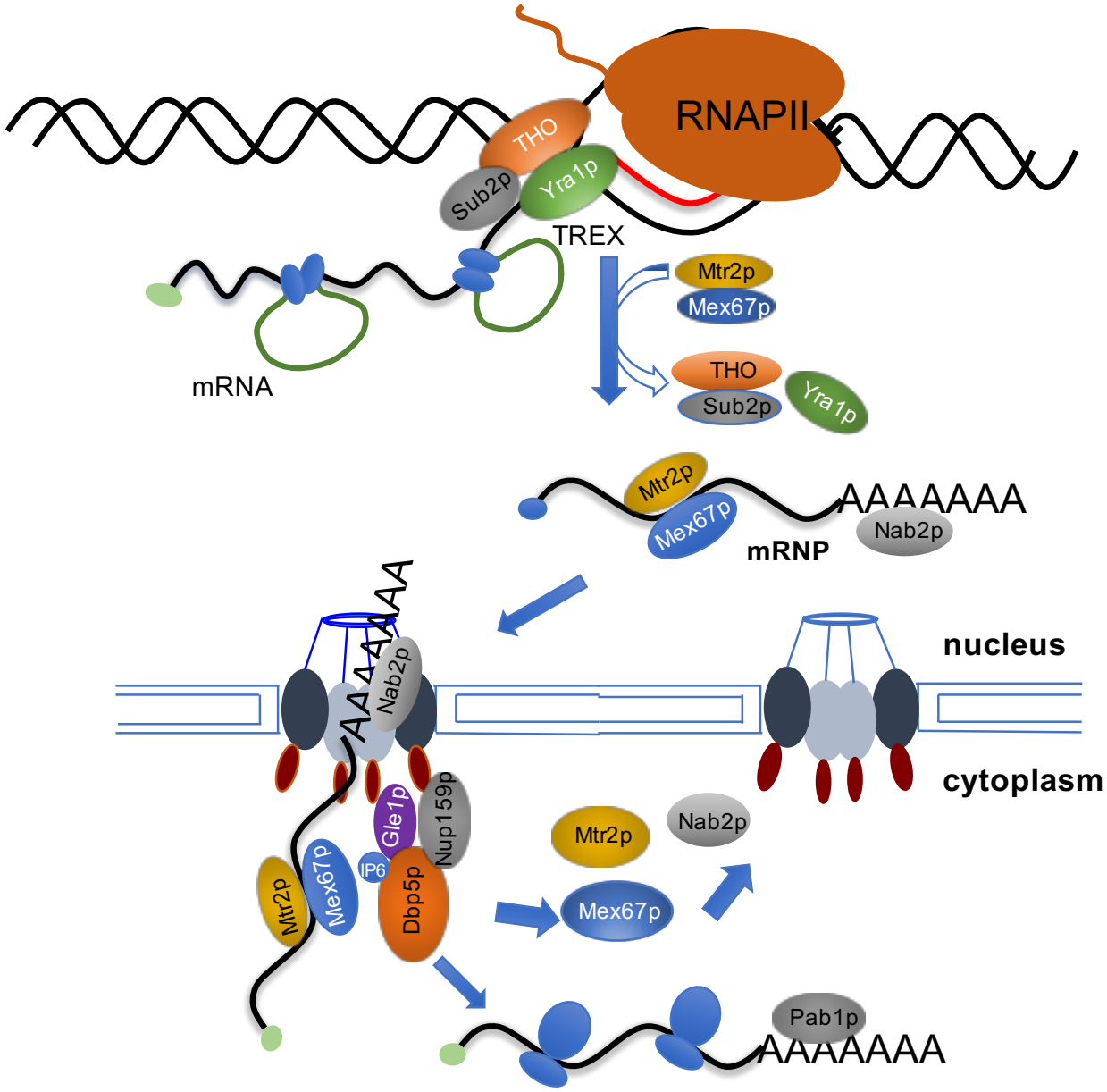


Figure 1-5. Schematic of mRNA export. During co-transcriptional processing, mRNAs become associated with the THO complex, which recruits adapter proteins such as Yra1p and Sub2p forming the TREX complex. Binding of the Mex67p-Mtr2p heterodimer leads to release of the adapter protein Yra1p to form an export competent mRNP that can be exported through NPCs. On the cytoplasmic side of the NPC, the mRNP is engaged by Dbp5p to promote removal of the Mex67p-Mtr2p heterodimer from the mRNP, thus releasing the mRNP into the cytoplasm. Dbp5p is stimulated by NPC associated proteins Nup159p and Gle1p with the small molecule inositol hexakisphosphate (IP_6). The nuclear poly(A) binding protein Nab2p is also replaced by the cytoplasmic poly(A) binding protein Pab1p following export.

1.5 Regulation of mRNP export

It is known that gene expression can be regulated post-transcriptionally at the level of mRNA export. For example, a genome wide study from the Silver laboratory (Harvard University) showed that only a subset of transcripts are associated with Yra1p (Hieronymus and Silver, 2003), suggesting that there may be multiple distinct mRNA export pathways that could be differentially regulated to control gene expression. It has also been demonstrated that Npl3p and Nab2p facilitate distinct export pathways (Green *et al.*, 2002; Hector *et al.*, 2002; Vinciguerra *et al.*, 2005). Specifically, these RBPs act with different ubiquitin E3 ligases, Tom1p and Rsp5p, which modify each RBP to regulate their function in mRNA export (Duncan *et al.*, 2000; Gupta *et al.*, 2007). For example, mutations in Tom1p blocks export of Nab2p associated mRNAs to the cytoplasm without impacting transcripts that are associated with Npl3p (Iglesias *et al.*, 2010). These data suggest that different mRNA export pathways exist, but we currently know little of how these pathways are directed to act on specific transcripts. Similarly, during a stress response bulk mRNA export is blocked, yet stress response mRNAs are still exported to the cytoplasm for protein synthesis (Saavedra *et al.*, 1996; Izawa, 2010). How the differential export of mRNAs is achieved is not well understood, but a more recent study showed that during stress, Mex67p is loaded directly onto mRNPs without adapter proteins in order to expedite the export of stress response mRNAs at the cost of export quality control (Zander *et al.*, 2016). Based on these multiple observations, it is clear that there are different pathways coupling mRNA processing and export, which could be differentially targeted to regulate gene expression. Given the scarcity of information about regulation of mRNA export, it is important that these pathways, and factors function within such pathways, are identified and further characterized.

1.6 Poly(A)-binding proteins and mRNA processing

There are two essential PABPs in *S. cerevisiae,* Pab1p and Nab2p. Generally, these proteins are thought to protect mRNAs from 3′ exonuclease mediated degradation, regulate transcription termination, and promote mRNA export (Hector *et al.*, 2002; Viphakone *et al.*, 2008). Both of these proteins shuttle between the nucleus and cytoplasm, however under steady-

state growth conditions Pab1p is predominantly localized in the cytoplasm, and conversely, Nab2p is localized in the nucleus (Green *et al.*, 2002; Brune *et al.*, 2005). Consistent with Nab2p being the major nuclear PABP, Nab2p associates with nascent polyadenylated mRNA and is exported to the cytoplasm with the mRNA, and is then replaced by Pab1p, allowing Nab2p to be reimported into the nucleus by Kap104p (Hector *et al.*, 2002; Dheur *et al.*, 2005; Fasken *et al.*, 2008). Nab2p is thought to function in mRNA export through its role in binding and recruiting specific factors (e.g. Mex67p) that promote export (Hector *et al.*, 2002; Fasken *et al.*, 2008; Iglesias *et al.*, 2010). Pab1p also appears to have some overlapping nuclear functions with Nab2p in the nucleus, as both PABPs can promote the regulation of correct poly(A) tail length *in vitro,* and mutation of either PABP leads to hyperadenylation *in vivo* (Hector *et al.*, 2002; Brune *et al.*, 2005; Dheur *et al.*, 2005; Viphakone *et al.*, 2008). Pab1p itself has also been linked to mRNA export (Brune *et al.*, 2005). In support of overlapping functions, studies mapping Pab1p and Nab2p binding across mRNAs have found that they bind a large set of shared substrates *in vivo* (Schmid *et al.*, 2012; Baejen *et al.*, 2014). Additionally, Pab1p binding does not only protect the mRNA from 3′ exonuclease degradation in the cytoplasm, but also functions with a cytoplasmic cap binding protein, eIF4G, to facilitate translation initiation by enhancing recruitment of the 40S ribosome (Wells *et al.*, 1998; Weill *et al.*, 2012).

Interestingly, in addition to protecting mRNAs from degradation, Nab2p is also involved in the decay of non-functional RNA substrates in the nucleus (Tuck and Tollervey, 2013). The dual roles of this poly(A) binding protein raises the issue of balancing the function of Nab2p between these two processes to ensure maintenance of nuclear RNA homeostasis. Indeed, a recent study from the Jensen lab (Aarhus University) showed that the nuclear accumulation of mRNAs has the ability to limit Nab2p activity and alter RNA processing in a manner similar to having a loss-of-function mutation in Nab2p (Tudek *et al.*, 2018b). However, this study was limited to the accumulation mRNAs, and it was not shown whether enrichment of any poly(A)-RNA substrate has the same ability to disrupt Nab2p function. In work presented in chapter III and IV, it is shown that the accumulation of polyadenylated ncRNAs has a similar impact on Nap2p function. This further supports the concept of nuclear RNA homeostasis, as these works together indicate that alterations in mRNA export, RNA decay, or ncRNA processing can impact nuclear poly(A)-RNA abundance, the availability and function of PABPs, and nuclear RNA processing.

1.7 mRNA quality control and surveillance mechanisms in the nucleus

As introduced and discussed in the proceeding sections, mature export competent mRNPs are produced through interactions of the pre-mRNA with various RBPs that facilitate nuclear processing steps. However, due to the number and complexity of steps required to generate the mature mRNP, errors do occur. To ensure that these errors do not lead to the production of aberrant mRNPs and altered gene expression, nuclear mRNPs are surveyed by nuclear quality control machineries. Defects in the 5′ capping, splicing, 3′ end formation, or polyadenylation of an mRNA are potential reasons why an mRNA would be targeted for degradation by the surveillance machinery.

1.7.1 Nuclear quality control of pre-mRNAs

Many studies on mRNA quality control mechanisms have focused on nonsense-mediated decay (NMD), which is a process in which mRNAs with pre-mature stop codons resulting from defective splicing, are degraded in the cytoplasm (Brogna and Wen, 2009; Schweingruber *et al.*, 2013). However, mRNAs are also subject to quality control processes in the nucleus prior to export. Quality control mechanisms here are not as well understood as NMD, and do not appear to be linked to any specific processing error; rather, targeting of an mRNA for degradation is proposed to be dependent on the time an mRNP spends in the nucleus (Porrua and Libri, 2013). This type of mechanism has the advantage of not requiring complex systems dedicated to recognizing different processing defects. Simply, the longer an mRNP remains in the nucleus, the greater the probability is that the mRNA is decayed. In support of this competition or timer model, tiling microarray data suggests that about half of yeast pre-mRNAs are degraded by the nuclear exosome (Gudipati *et al.*, 2012), which is higher than expected, and proposed to be the result of the longer times taken to generate a spliced mRNA.

Mechanistically, it has been suggested that defective mRNAs (e.g. retained introns) would have lower affinity for mRNA export factors, which would decrease their likelihood of export, lead to nuclear retention, and degradation due to their persistence in the nucleus. (Soheilypour and Mofrad, 2016). This concept is supported by data showing that the serine/arginine rich (SR)

RBP Npl3p present in an un-spliced mRNP favors binding of RNA decay machinery, but following splicing, Npl3p preferentially binds to the Mex67p-Mtr2p heterodimer to facilitate mRNP export (Gilbert and Guthrie, 2004; Kress *et al.*, 2008; Hackmann *et al.*, 2014; Zander *et al.*, 2016). Given that Npl3p is bound co-transcriptionally and is recruited to normally processed transcripts, it would function to promote either outcome based on splicing state; hence slow or defective splicing would lead to decay without a need to recognize an actual defect in the mRNP (Figure 1-6). A related mode of splicing quality control also involves the retention of faulty pre-mRNAs in the nucleus to promote decay through interactions between mRNP export adapters and the NPC basket components Mlp1p, Mlp2p, and Pml39p via the 5′ splice site of the mRNA (Dziembowski *et al.*, 2004; Galy *et al.*, 2004). In the context of 3′-end processing or polyadenylation defects, it is also observed that transcripts are held in the nucleus, at or near transcription sites, and undergo degradation by nuclear RNA decay machinery (Hilleren *et al.*, 2001). These data all are in line with the concept that nuclear retention is a mechanism to promote mRNA surveillance and decay. Under this mode of quality control, efficient recognition and decay of transcripts by RNA decay machinery is critical, since failure to decay defective transcripts would allow mRNAs to escape the nucleus to continue in the gene expression program or result in a buildup of decay substrates that could deplete cellular pools of RBPs. Of the various proteins and protein complexes with nuclease activity, the exosome is the major complex functioning within the nucleus in RNA surveillance and decay, in addition to having roles in various other aspects of RNA biogenesis (Lykke-Andersen *et al.*, 2009; Chlebowski *et al.*, 2013; Schneider and Tollervey, 2013).

1.7.2 The RNA exosome

The exosome functions within the cytoplasm and nucleus to facilitate RNA biogenesis, surveillance, and decay. The exosome core consists of six proteins that form a barrel structure with a central channel, upon which sits a three protein subunit cap (Figure 1-7) (Chlebowski *et al.*, 2013; Schneider and Tollervey, 2013). The central channel of the exosome is formed by the RNase phosphorolysis (PH) domain containing subunits (Rrp41p, Rrp42p, Rrp43p, Rrp45p, Rrp46p, and Mtr3p), while the cap contains Rrp4p, Rrp40p and Csl4p each harbouring S1

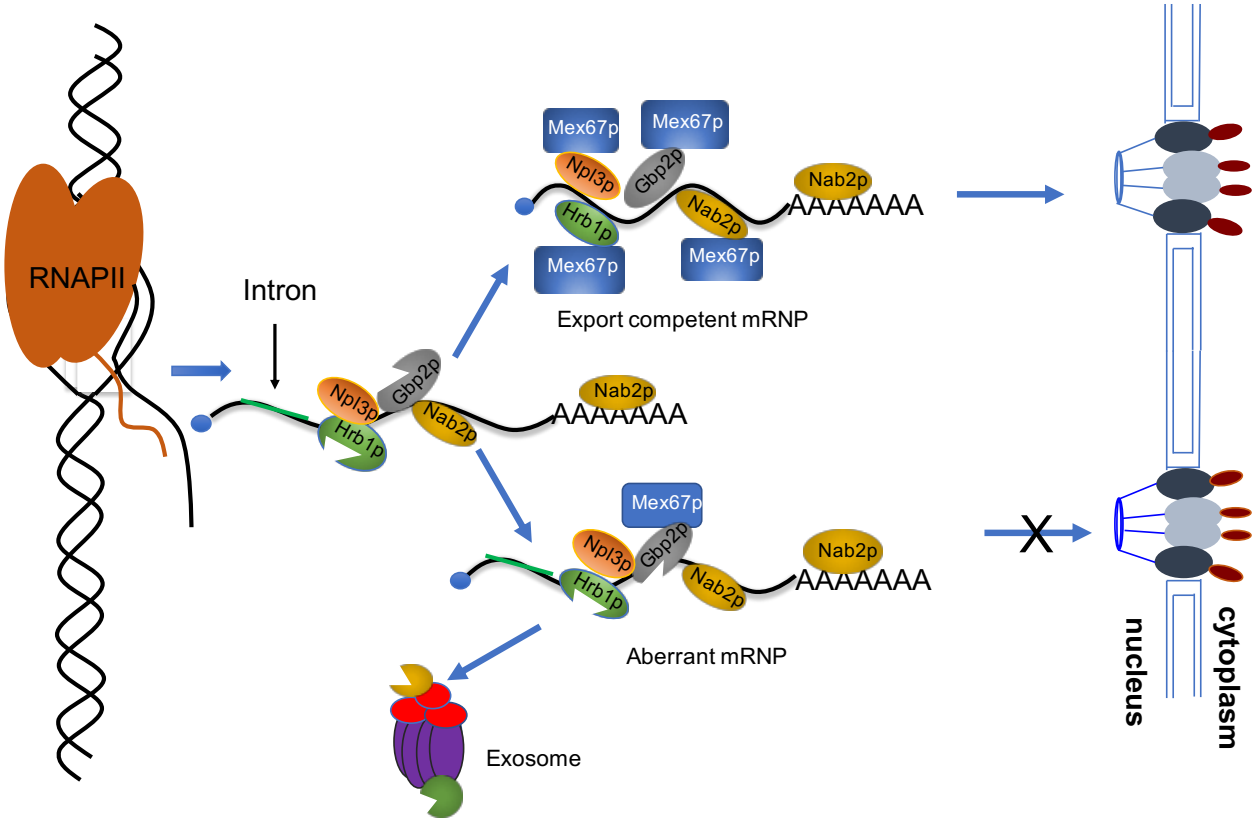


Figure 1-6. Quality control of mRNA splicing. Adapter proteins (Nab2p - yellow, Hrb1p - green, Gbp2p - gray, and Npl3p - orange) recruit downstream effector proteins that include the Mex67p-Mtr2p heterodimer to the mature mRNP for export through nuclear pore complexes. In the context of an un-spliced transcript with an intron (green), conformations of Hrb1p and Gbp2p favor recruitment of RNA decay machinery over the Mex67-Mtr2p export adaptor.

RNA-binding domains. The S1 domain is named after the ribosomal protein S1, where the domain was first identified, which was then later identified in various other RNA binding proteins (Boni *et al.*, 1991). Rrp4p and Rrp40p contain K homology(KH) domain, which were first identified in the heterogeneous nuclear ribonucleoprotein K (Siomi *et al.*, 1993). All nine components of the exosome are collectively known as Exo-9, a complex that lacks nuclease activity, but is still essential for yeast viability (Liu *et al.*, 2006). The nuclear and cytoplasmic Exo-9 complex associates with the catalytic subunit Rrp44p/Dis3p, an essential multi-domain protein containing both exonuclease and endonuclease activities (Lebreton *et al.*, 2008; Bonneau *et al.*, 2009; Schaeffer *et al.*, 2009). Dis3p specifically binds to the exosome complex through interactions with Rrp41p, Rrp43p, and Rrp45p (Schneider *et al.*, 2009). In addition to Dis3p, the nuclear exosome contains a non-essential catalytic subunit with distributive 3′-5′ exonuclease activity known as Rrp6p (Briggs *et al.*, 1998; Burkard and Butler, 2000). Distributive exonuclease activity is defined as a process where the enzyme dissociates from the RNA substrate after cleaving a single nucleotide from the RNA. While Dis3p associates with the bottom of the exosome barrel, Rrp6p interacts with cap components (Figure 1-7) (Stead *et al.*, 2007).

As the catalytic components of the exosome, Rrp6p and Dis3p participate in mRNP surveillance, rRNA, snoRNA, and snRNA biogenesis, and the decay of pervasive transcripts (Lykke-Andersen *et al.*, 2009). There are also non-exosome associated activities of Rrp6p, including the processing of rRNAs (Callahan and Butler, 2008). In addition, the exosome has roles in RNA decay mediated gene regulation, including the decay of meiotic mRNAs during the mitotic cell cycle in *S. cerevisiae* (Harigaya *et al.*, 2006), and Rrp6p-dependent decay of histone transcripts after DNA replication (Canavan and Bond, 2007; Reis and Campbell, 2007). The mechanisms by which the exosome targets these diverse substrates for degradation or processing remains unclear. However, it is well known that the function of the exosome, to a large extent, depends on associated protein complexes, which include the Trf4p-Air2p-Mtr4p polyadenylation (TRAMP) and Nrd1p-Nab3p-Sen1p (NNS) complexes (see sections 1.7.4 & 1.7.5).

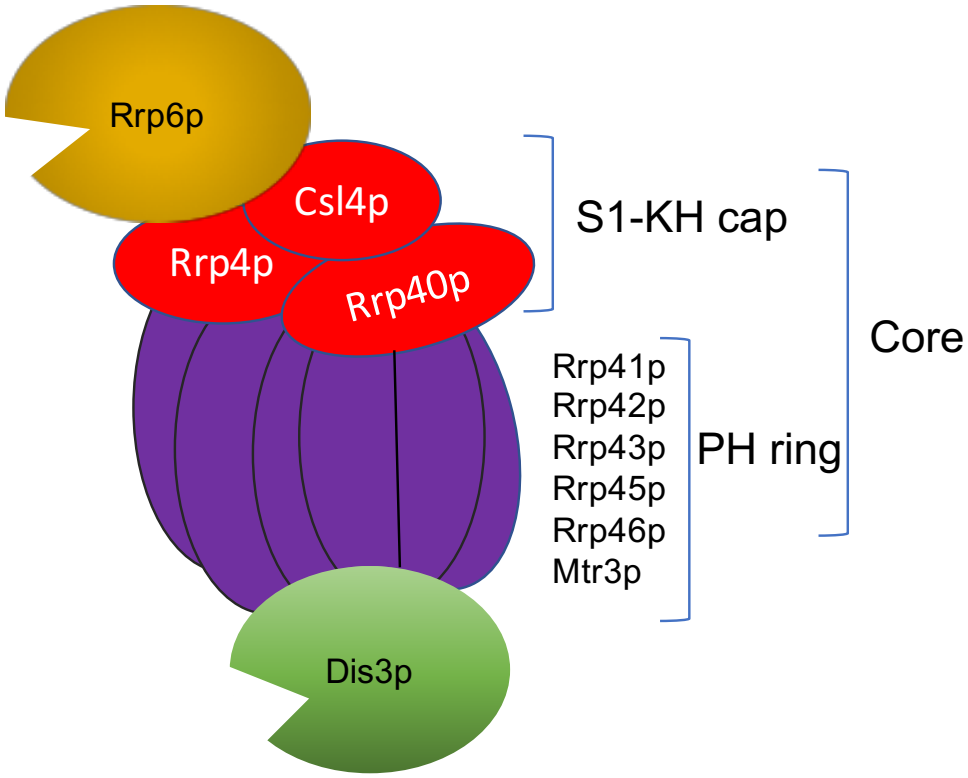


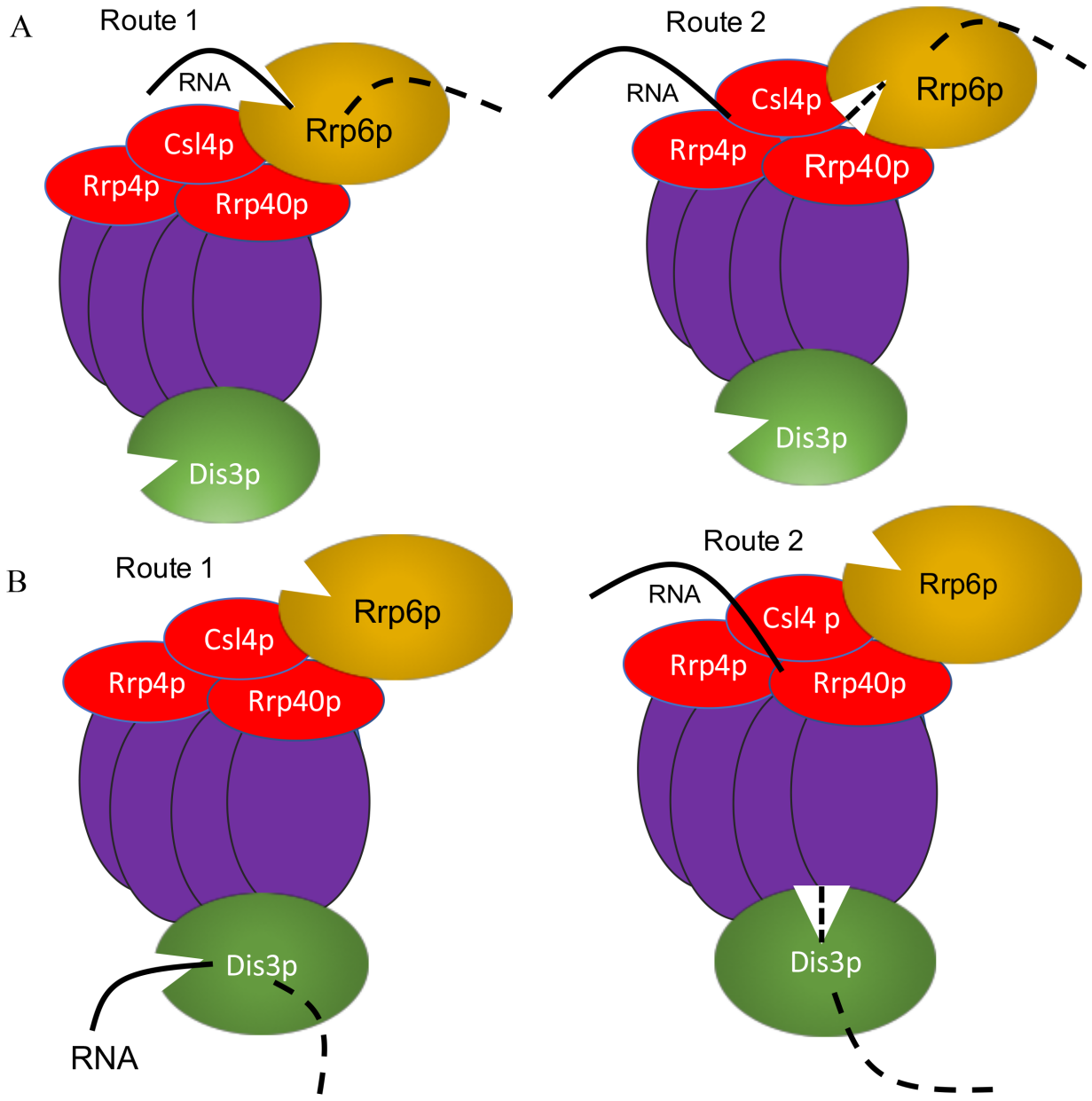
Figure 1-7. Schematic of the RNA exosome. The exosome consists of a nine-subunit core (purple and red) and two catalytic subunits, Dis3p (green) and Rrp6p (yellow). The core of the exosome can be divided into two parts, the central channel and the cap like structure. The central channel of the exosome is formed by six RNase phosphorolysis (PH) containing protein subunits (purple)(i.e. Rrp41p, Rrp42p, Rrp43p, Rrp45p, Rrp46p, Mtr3p). The cap of the exosome is formed by three proteins containing S1 and KH RNA-binding domains (red) (i.e. Csl4p, Rrp40p, Rrp4p).

1.7.3 Exosome mediated RNA degradation

Data from X-ray crystallography studies detailing Rrp6p and the exosome complex suggest that RNAs can be degraded through directly engaging Rrp6p without the aid of any core exosome components (Makino *et al.*, 2015) (Figure 1-8A left). In an alternative mechanism, RNAs may be threaded into the active site of Rrp6p from the side, between the cap and central channel of Exo-9 (Wasmuth and Lima, 2012; Wasmuth *et al.*, 2014) (Figure 1-8A right). Structural studies have also demonstrated that Dis3p degrades RNA via two distinct pathways. Route one involves degradation of RNA substrates, such as the 5S rRNA, by directly binding Dis3p (Kadaba *et al.*, 2006; Dziembowski *et al.*, 2007; Han and van Hoof, 2016) (Figure 1-8B left). Route two involves passing an RNA through the channel of the exosome in a 3′-5′ direction to be degraded by the exonuclease activity of Dis3p (Malet *et al.*, 2010; Liu *et al.*, 2014) (Figure 1-8B right). The feeding of an RNA through the Exo-9 channel to reach the active site of Dis3p requires ~30 nucleotides of unstructured RNA (Briggs *et al.*, 1998). Dis3p has been shown to degrade RNA substrates in isolation, but the nuclease activity of Dis3p is increased to a large extent after addition of Exo9 (Wasmuth and Lima, 2012). Recently, Wasmuth *et. al*., also demonstrated that Rrp6p can allosterically stimulate the nuclease activity of Dis3p by widening the path through the central channel to facilitate access to Dis3p (Wasmuth *et al.*, 2014). Thus, Exo9 is able to facilitate both the recognition and degradation of RNA substrates by Rrp6p and Dis3p.

1.7.4 The Trf4p-Air2p-Mtr4p polyadenylation (TRAMP) complex

The exosome can degrade an RNA via channeling it through the core of the exosome if the RNA is unstructured and can reach Dis3p (Figure 1-8B right), which requires ~30 nucleotides of unstructured RNA (Bonneau *et al.*, 2009). Given that many RNAs do not possess such a long unstructured region, nuclear RNA degradation by this pathway is aided by the TRAMP complex (Figure 1-9). The main functions of the TRAMP complex are to add a short poly(A) tail to RNA substrates that need to be targeted by the exosome and to aid in the unwinding of structured RNAs through the activity of the RNA helicase Mtr4p (Hardwick and Luisi, 2013). The combined addition of a poly(A) tail and unwinding of secondary structure together allow RNA



Figures 1-8. Exosome mediated RNA degradation. (A) RNAs can be degraded by the nuclear specific ribonuclease Rrp6p (blue) through two different means: (1) an RNA is recognized and degraded by the activity of Rrp6p directly without the core exosome components (red and purple) or (2) protein subunits that make up the exosome cap (red) recognize a substrate and aid in delivering the RNA to Rrp6p for degradation. (B) Dis3p can engage an RNA for decay via: (1) direct binding of the RNA to the nuclease or (2) channeling the RNA through the core of the exosome in order to reach the active site of Dis3p.

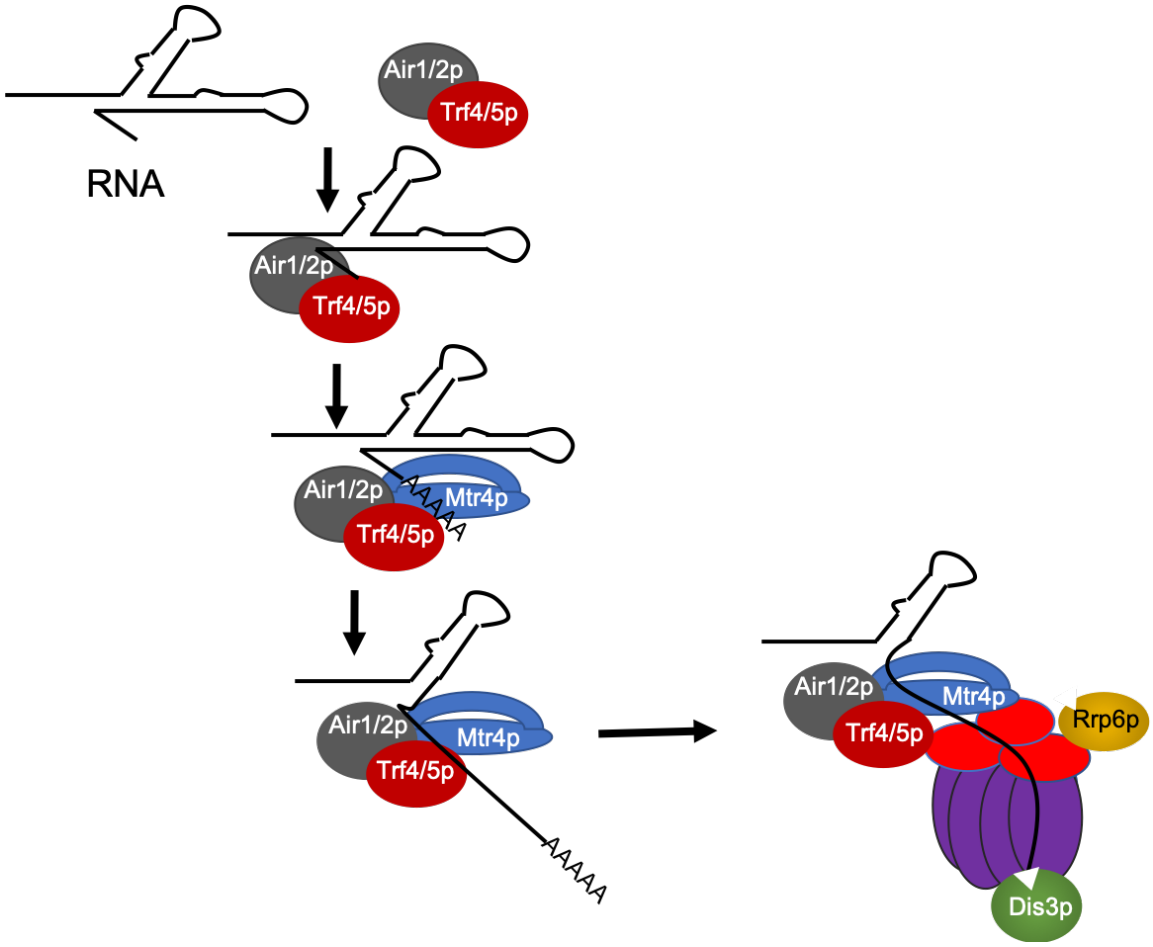


Figure 1-9. The TRAMP complex and RNA degradation. RNA substrates associate with the TRAMP complex via RNA binding proteins Air1p/Air2p (grey) to bring the associated poly(A) polymerase, either Trf4p or Trf5p (red), to the RNA. Trf4p/Trf5p mediated polyadenylation of the RNA substrate, plus unwinding of RNA secondary structure via the activity of Mtr4p, allow the RNA substrate to be channelled to Dis3p for degradation.

substrates to be passed through the central channel of the exosome to be engaged by Dis3p (LaCava *et al.*, 2005). Similarly, in the cytoplasm the super killer (SKI) complex is essential for the unwinding of RNA secondary structure to facilitate passage of the RNA through the core of the exosome (Halbach *et al.*, 2013). As part of the TRAMP complex, the non-canonical poly(A)-polymerase protein subunit, either Trf4p or Trf5p, is responsible for polyadenylating target RNAs for processing by the exosome (Vanácová *et al.*, 2005; Egecioglu *et al.*, 2006; Houseley and Tollervey, 2006). Given the presence of Trf4p or Trf5p in the complex, the complex is referred to as TRAMP4 or TRAMP5 to distinguish between these two possibilities. In contrast to Pap1p, the poly(A)-polymerase active on most mRNAs, Trf4p and Trf5p lack an RNA-binding domain, hence RNA binding requires other subunits of the TRAMP complex, namely Air1p/Air2p (LaCava *et al.*, 2005; Vanácová *et al.*, 2005; Wyers *et al.*, 2005). Mtr4p, the other component of the TRAMP complex, in addition to promoting RNA unwinding through a helicase activity as mentioned above, is composed of other protein interaction domains (Jackson *et al.*, 2010; Taylor *et al.*, 2014). These domains are known to bind Nop53p and Utp18p to target rRNAs for processing by the exosome (Tudek *et al.*, 2014; Thoms *et al.*, 2015). Therefore, in addition to the TRAMP complex aiding the exosome in degrading target RNAs, this complex also engages other protein partners to facilitate substrate recognition.

1.7.5 The Nrd1p-Nab3p-Sen1p (NNS) Complex

Exosome activity is further directed by the Nrd1p-Nab3p-Sen1p (NNS) complex. The NNS complex is involved in nuclear RNA quality control and transcription termination of mRNAs, snoRNAs, and other short RNAs (Vasiljeva and Buratowski, 2006). During transcription, the NNS complex is recruited to a nascent RNA via Nrd1p binding to the phosphorylated CTD of RNAP II (Conrad *et al.*, 2000; Steinmetz *et al.*, 2001; Vasiljeva and Buratowski, 2006). In addition to interacting with the RNAPII CTD, Nrd1p binds to a consensus sequence in nascent RNAs using an RNA recognition motif (RRM), as does Nab3p (Carroll *et al.*, 2007). Following recruitment to the RNA, the RNA helicase Sen1p facilitates termination and release of RNAPII transcripts through ATP hydrolysis (Hazelbaker *et al.*, 2013) (Figure 1-10). The domain of Nrd1p that is involved in binding to the CTD of RNAPII also binds Trf4p of

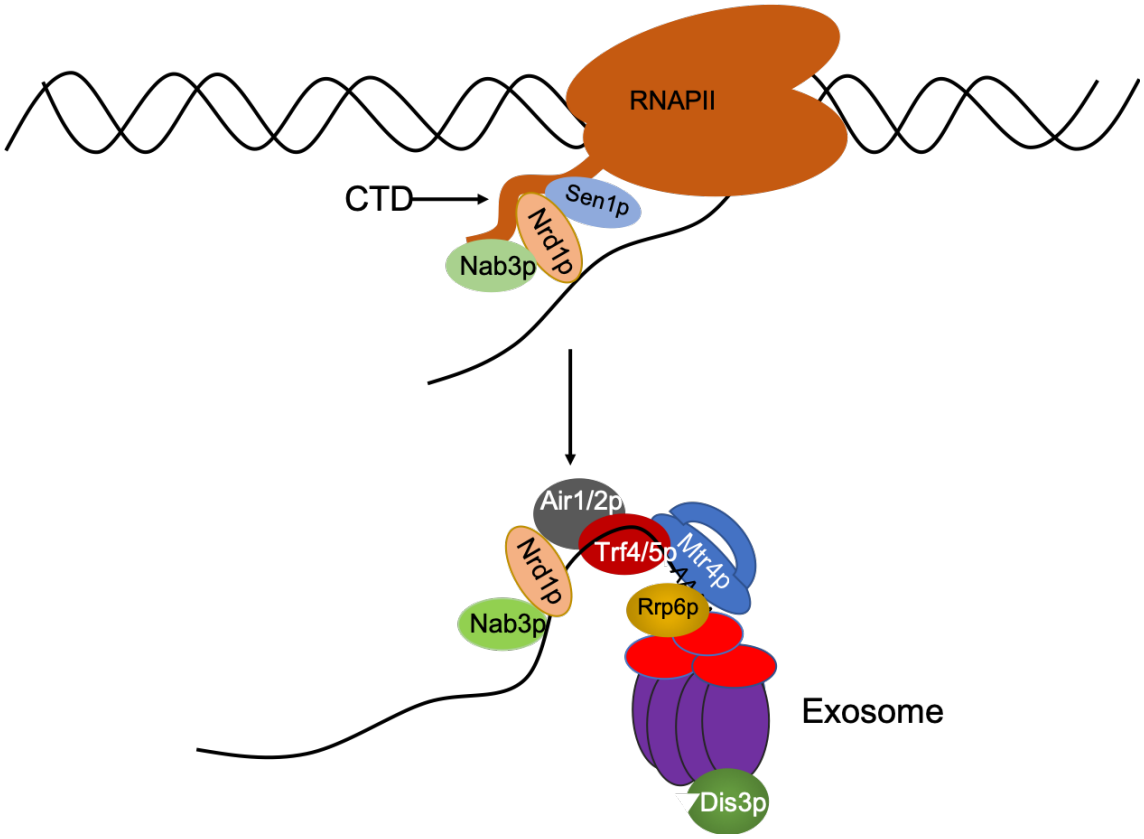


Figure 1-10. The NNS complex mediated RNA degradation by the exosome. ncRNAs transcribed by RNAPII are terminated by the interaction between the CTD of RNAPII and the NNS complex. Nrd1p (pink) and Nab3p (green) interact with nascent RNAs and recruit an additional component of the NNS complex, Sen1p, which facilitates transcription termination. Subsequently, the NNS complex recruits the TRAMP complex, Trf4p/Trf5p (red), which add a short poly(A) tail to ncRNAs and is delivered to the exosome for degradation by Mtr4p. The exosome is recruited to the ncRNA through the interaction between the TRAMP complex and the exosome core for either processing or decay.

the TRAMP complex, and provides a mechanism to promote exosome activity on the targeted RNA (Arndt and Reines, 2015). Recruitment of the TRAMP complex provides poly(A)-polymerase activity for the addition of a short poly(A) tail to promote exosome recognition and degradation (LaCava *et al.*, 2005; Vanácová *et al.*, 2005; Wyers *et al.*, 2005; Houseley and Tollervey, 2009). RNA-binding data generated from UV-crosslinking approaches have demonstrated that in addition to RNAPII transcripts, the NNS complex is also involved in the processing of RNAPI and RNAPIII transcripts in association the TRAMP complex (Webb *et al.*, 2014; van Nues *et al.*, 2017). Interestingly, the *NRD1* mRNA is self-regulated by NNS directed degradation via multiple Nrd1p binding sites, which feeds back to regulate Nrd1p abundance and NNS complex activity (Arigo *et al.*, 2006a).

1.7.6 The exosome cofactors Rrp47p an Mpp6p

In addition to the TRAMP and NNS complexes, Rrp47p is another co-factor of the nuclear exosome that is required for Rrp6p mediated processing of stable ncRNAs, such as rRNAs, snRNAs, snoRNAs, and degradation of cryptic unstable transcripts (CUTs) (Mitchell *et al.*, 2003; Arigo *et al.*, 2006b). CUTs are unstable noncoding RNAs that are transcribed by RNAPII and efficiently degraded by surveillance machineries, as such they are only robustly detected in the background of RNA decay mutants (Wyers *et al.*, 2005). In addition to the degradation of CUTs and aberrant rRNAs (Stead *et al.*, 2007), Rrp47p is needed for both core-dependent and core-independent processing of stable ncRNAs by Rrp6p (Mitchell *et al.*, 2003). Data from structural studies have demonstrated that the interaction of Rrp6p with Rrp47p provides a binding surface important for Mtr4p association with the nuclear exosome (Schuch *et al.*, 2014). Similarly, Mpp6p is an RNA binding protein that functions to support 5.8S rRNA processing, pre-mRNA processing, and the degradation of pervasive transcripts (Schilders *et al.*, 2005; Milligan *et al.*, 2008). Deletion of *MPP6* has been reported to be synthetic lethal with the deletion of *RRP6* or *RRP47,* and generally has phenotypes similar to mutants that alter TRAMP activity (Falk *et al.*, 2017a). In line with this, a large number of the transcripts stabilized upon deletion of *MPP6* are also stabilized in a *dis3-1* mutant strain (Milligan *et al.*, 2008). Recent work also supports a role for Mpp6p in facilitating interactions between Mtr4p and Rrp40p of the

core exosome (Falk *et al.*, 2017a), which is enhanced in the presence of Rrp47p (Stuparevic *et al.*, 2013). Overall, it is apparent that these additional exosome co-factors facilitate exosome recognition of RNA substrates for their efficient processing or decay.

1.8 RNA Polymerase II (RNAPII) transcription of noncoding RNAs

1.8.1 snoRNA transcription and processing

A major class of functional RNAPII transcribed ncRNAs are snoRNAs. snoRNAs are short, non-polyadenylated ncRNAs that function in the nucleolus (Kiss, 2002). snoRNAs can be divided into two classes, Box C/D and H/ACA snoRNAs, based on conserved sequences. Most snoRNAs also contain sequences complementary to the rRNA and snRNA targets they engage to direct 2-O′ methylation (Box C/D) and pseudouridylation (Box H/ACA) (Kiss, 2002). A few snoRNAs also function in the processing of rRNAs (Watkins and Bohnsack, 2012). To perform these functions, snoRNAs associate with various protein factors to form snoRNP complexes (Tran *et al.*, 2003). The protein components of the snoRNP provide enzymatic activity, while snoRNAs provide target specificity. In addition to rRNA processing, snoRNAs have also been linked to mRNA splicing, chromatin maintenance, and RNA editing (Wu *et al.*, 2016). In mammals, snoRNA genes are located in polycistronic transcription units; however, in *S. cerevisiae* most snoRNA genes are individual transcription units (Piekna-Przybylska *et al.*, 2007), which are transcribed as longer RNAs that require trimming of both the 5′ and 3′ end to produce a mature snoRNA (Allmang *et al.*, 1999a; van Hoof *et al.*, 2000a). Specifically in the case of snoRNA 3′ end processing, the NNS, TRAMP, and exosome complexes are involved (Kufel and Grzechnik, 2019). Briefly, upon transcription of Nrd1p and Nab3p binding sites in a snoRNA, binding of NNS components slows down RNAPII leading to Sen1p dependent termination of transcription (Porrua *et al.*, 2016). After termination by the NNS complex, pre-snoRNAs are transiently polyadenylated by the TRAMP complex, which leads to the subsequent recruitment of the exosome and/or Rrp6p for 3′→ 5′ trimming of snoRNAs (Tudek *et al.*, 2014) (Figure 1-11). Consequently, in the absence of exosome or Rrp6p function, 3′ extended snoRNA pre-cursors accumulate (van Hoof *et al.*, 2000a).

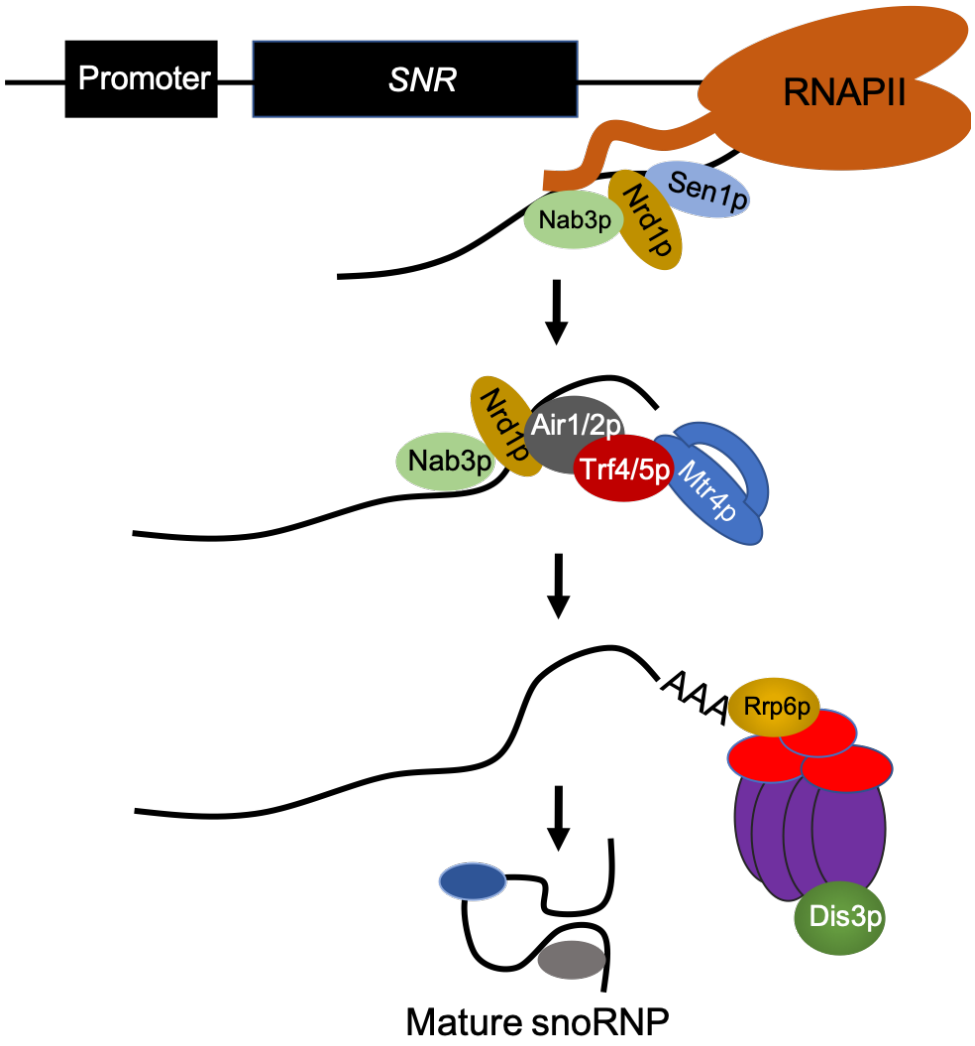


Figure 1-11. Exosome mediated snoRNA processing. snoRNAs are transcribed as long pre-snoRNAs, which undergo processing by the NNS, TRAMP, and exosome complexes. Binding of Nrd1p (yellow) and Nab3p (blue) to 3′ end of pre-snoRNAs recruits Sen1p (light blue), which terminates snoRNA transcription. Following termination, pre-snoRNAs are recognized by Air1p/Air2p, transiently polyadenylated by Trf4p/Trf5p, delivered to the exosome with the help of Mtr4p for 3′ processing by the nuclear exosome.

1.8.2 Pervasive transcription by RNA Polymerase II

Pervasive transcription refers to the production of RNAPII transcripts that do not encode a protein and have no established cellular function (Wyers *et al.*, 2005). Pervasive transcripts are defined by the term "pervasive" given the widespread origins of the transcripts from most areas of the genome. In some ways, this can be thought of as transcriptional noise; however, in some instances specific cellular functions may have also evolved for individual pervasive transcripts (Martens *et al.*, 2004; Camblong *et al.*, 2007; Houseley *et al.*, 2008). One of the first classes of unconventional RNAPII transcripts discovered in yeast was upon mutation of the chromatin remodeler Spt6p, which plays a key role in the correct nucleosome positioning to prevent aberrant transcription from cryptic promoters (Kaplan *et al.*, 2003). Subsequently, transcriptome analysis of an *rrp6Δ* strain revealed normally hidden transcripts termed cryptic unstable transcripts (CUTs), which were discovered to be constantly produced but rapidly degraded by the nuclear exosome (Davis and Ares, 2006; Houalla *et al.*, 2006). A recent study by Vera and Dowell have further defined CUTs using RNA-seq data acquired from a *rrp6Δ* strain, providing a more accurate annotation of CUTs in yeast (Vera and Dowell, 2016). Other studies have also defined additional classes of pervasive transcripts termed stable unannotated transcripts (SUTs) (Xu *et al.*, 2009) and Xrn1p sensitive unstable transcripts (XUTs) (Van Dijk *et al.*, 2011). SUTs are a group of transcripts that are present in wild type cells at a detectable level, but do not overlap with annotated genomic features or have defined functions (Xu *et al.*, 2009). XUTs are longer polyadenylated transcripts that are exported to the cytoplasm and are decayed by the cytoplasmic 5′-3′ exoribonuclease 1, Xrn1p, and are therefore stabilized in Xrn1p mutants (Van Dijk *et al.*, 2011).

Mechanisms for termination and decay of these various groups of pervasive transcripts are distinct; however, they often share the same transcription initiation site. For example, many of these transcripts originate from nucleosome depleted regions (NDRs) at either the 5′ or 3′ end of a gene as revealed by native elongating transcript sequencing (NET-seq) analysis (Churchman and Weissman, 2011, 2012). Similarly, CUTs and SUTs are often produced by divergent transcription at the promoter region of a gene, suggesting gene promoters are intrinsically bidirectional (Xu *et al.*, 2009). Given that pervasive transcripts are generated from promoter regions, the production of some pervasive transcripts appears to be regulated by the binding of

transcription factors in response to environmental stimuli (Lardenois *et al.*, 2011; van Werven *et al.*, 2012). For example, increased expression of "meiotic-specific noncoding transcripts" (MUTs) occurs in response to nutrient deprivation and during the entry of diploid cells into meiosis (Lardenois *et al.*, 2011). Given the level of pervasive transcription, and association of these transcripts with much of the same machinery as mRNAs and ncRNAs, the inappropriate stabilization of pervasive transcripts could be a situation where the balance between RNA and RBPs is disturbed and nuclear RNA homeostasis is impacted.

1.8.3 The function and regulation of pervasive transcription

To date, only a few pervasive transcripts have been shown to have a cellular function. In some instances, pervasive transcription is found to induce or decrease expression of a nearby genes. For example, expression of the *SRG1* ncRNA, which is upstream of *SER3*, can repress the transcription of the *SER3* gene in high serine conditions (Thebault *et al.*, 2011). Another example is found associated with the meiotic transcript *IME4*, in this case antisense transcripts overlap with the *IME4* promoter and inhibit transcription in haploid cells (Hongay *et al.*, 2006; van Werven *et al.*, 2012). Pervasive transcripts can also exert their function through chromatin modification, for example, an unstable *PHO84* antisense ncRNA stimulates histone deacetylation at the *PHO84* promoter and represses *PHO84* expression (Camblong *et al.*, 2007, 2009). A more recent effort to assign functions to pervasive transcripts was undertaken by deleting these transcriptional units from the genome followed by screening for growth phenotypes (Parker *et al.*, 2017, 2018). This study identified four essential SUTs, which were attributed to functions in mediating neighbouring gene expression (Parker *et al.*, 2018). As such, it is likely that most of these transcripts are non-functional and need to be cleared by RNA decay pathways to prevent their accumulation, which may be toxic to the cell.

Cellular levels of pervasive transcripts are generally kept in check by two different pathways. The first pathway is directed at controlling leaky transcription from NDRs by remodeling chromatin structure. For example, the Rpd3S histone deacetylase complex (HDAC) is known to act on histone H4 to prevent spurious transcription at intragenic regions (Carrozza *et al.*, 2005). Spurious transcription from NDRs is also prevented by the ATP-dependent chromatin

remodeler, Isw2p, by regulating nucleosome positioning within NDRs (Whitehouse *et al.*, 2007). The second mode of control acts at the level of nuclear RNA decay, which can be co-transcriptional within the nucleus, or occur post-transcriptionally within the nucleus or cytoplasm (Jensen *et al.*, 2013). For example, the NNS complex can terminate CUT transcription and promote decay by recruiting the TRAMP and exosome complexes to the CUT (Figure 1-12A) (Jensen *et al.*, 2013), as described above for snoRNAs. The interaction of RNAPII with the NNS complex for termination, as compared to the CPF complex, is one of the major determinants used to differentiate canonical mRNAs from CUTs and to target CUTs for decay. Use of the NNS vs. CPF complex for termination depends in part on the position of the terminator signal with respect to the transcription start site (Jacquier, 2009). Given this dependence, recent work has also defined a class of pervasive transcripts that are stabilized in mutants that disrupt NNS complex function, which are termed NNS unstable transcripts (NUTs) (Schulz *et al.*, 2013). However, there are other classes of pervasive transcripts that are terminated by CPF-CFI/II and require post-transcriptional decay (Figure 1.12B). For example, after transcription termination, XUTs are exported to the cytoplasm and decayed by Xrn1p, while SUTs undergo cleavage and polyadenylation, but are not bound by Mex67p, and are thus decayed in the nucleus (Tuck and Tollervey, 2013). The similarity between SUTs and mRNAs, including their interaction with a shared set of RBPs, may lead to competition between SUTs and mRNAs for RNA processing factors, which may be a limited set of resources. A paradigm of potential gene expression regulation that has not been deeply explored.

1.9 rRNA biogenesis

1.9.1 rRNA transcription by RNA Polymerase I (RNAPI)

Biogenesis of both the small and large ribosomal subunits begins with the production of the 18S, 5.8S, and 25S rRNAs from a single 35S transcript from a tandem repeat of ~150 rRNA gene copies in yeast (Thiry and Lafontaine, 2005). rRNA transcription and ribosome biogenesis occurs within the nucleolus, a non-membrane bound organelle organized around rDNA (Hernandez-Verdun, 2006; Staub *et al.*, 2006; Klinge and Woolford, 2019). RNAPI is

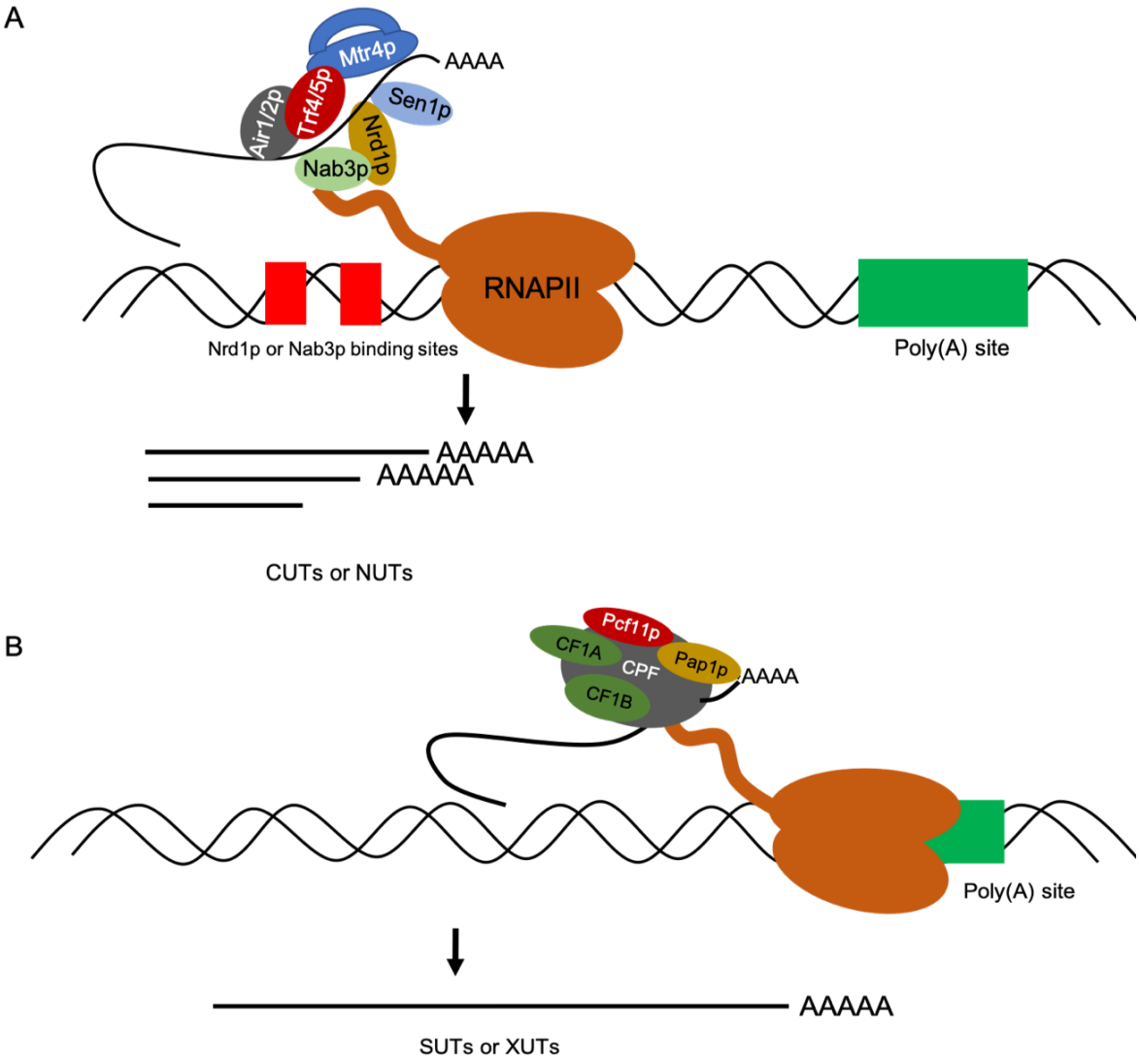


Figure 1-12. Pervasive transcription of CUTs, NUTs, SUTs, and XUTs. (**A**) Recruitment of the NNS complex occurs through the CTD of RNAPII and recognition of binding sites on the RNA by Nrd1p and Nab3p, which promote recruitment of the RNA helicase Sen1p. The ATPase activity of Sen1p facilitates the release of RNAPII leading to the termination of CUT or NUT transcription. The TRAMP (Trf4p/Trf5p, Air1p/Air2p, Mtr4p) complex polyadenylates NUTs for their degradation, while CUTs are not often polyadenylated and are shorter than NUTs. (**B**) Pervasive transcripts such as SUTs and XUTs are terminated by the CPF pathway, leading to polyadenylation of the transcript by Pap1p. Both SUTs and XUTs are often longer than CUTs and NUTs, with XUTs generally being exported and degraded in the cytoplasm by Xrn1p.

responsible for the synthesis of rRNA, which requires very high transcription rates to meet the cellular demand for functional ribosomal subunits to sustain protein synthesis (Schneider, 2012). RNAPI binds with rRNA genes through interactions with transcription factors, Rrn6p, Rrn7p, Rrn11p, that form a pre-initiation complex on DNA (Woolford and Baserga, 2013). RNAPI transcription also requires other factors, including Rrn3p and the heterodimeric core transcription factors (Moorefield *et al.*, 2000; Blattner *et al.*, 2011). Individual rRNA precursors generated by RNAPI transcription must undergo co-transcriptional and post-transcriptional processing to produce mature 18S, 5.8S, and 25S rRNAs (Henras *et al.*, 2015; Klinge and Woolford, 2019). Processing involves the removal of the external transcribed spacers (5′-ETS and 3′-ETS) sequences, as well as internal transcribed spacers (*ITS1* and *ITS2*), to generate the individual rRNAs. Cleavage of the 35S rRNA requires two 5′-3′ exonucleases (Rat1p and Xrn1p), the 3′-5′ exonuclease activity of the exosome complex, and two endonucleases (RNase MRP and Rtn1p) (Elela *et al.*, 1996; Fatica and Tollervey, 2002; Lindahl *et al.*, 2009). As part of this processing, pre-rRNA transcripts are bound by ribosomal proteins (RPs), ribosomal biogenesis factors, and snoRNAs to direct maturation and the folding of the rRNA into functional ribosomal subunits.

1.9.2 Maturation of the 18S rRNA

Maturation of the 18S rRNA involves liberation of the 18S transcript from the larger pre-rRNA by cleavages that remove the 5′-ETS and 3′-ITS1 sequences through endonucleolytic cleavage and trimming (Figure 1-13). This is accomplished co-transcriptionally by endonucleolytic cleavage at positions A0, A1, and A2 to generate a 20S precursor from the nascent rRNA transcript (Koš and Tollervey, 2010). In addition, rRNAs can be processed through an alternative pathway post-transcriptionally, which first involves cleavage at B0 in the 3′-ETS pre-rRNA to generate the 35S pre-rRNA (Kufel and Grzechnik, 2019). This is followed by cleavage at the A3 site to generate 23S and 27S pre-rRNAs, then cleavages within the 23S transcript (sites A0, A1, and A2) to generate the 20S rRNA (Koš and Tollervey, 2010). The 20S pre-rRNA produced through either co-transcriptional or post-transcriptional processing associates with pre-ribosomal proteins. The formed pre-ribosomal subunit is then exported to the cytoplasm where final cleavage occurs at the D site to produce the mature 18S rRNA, which is part of the

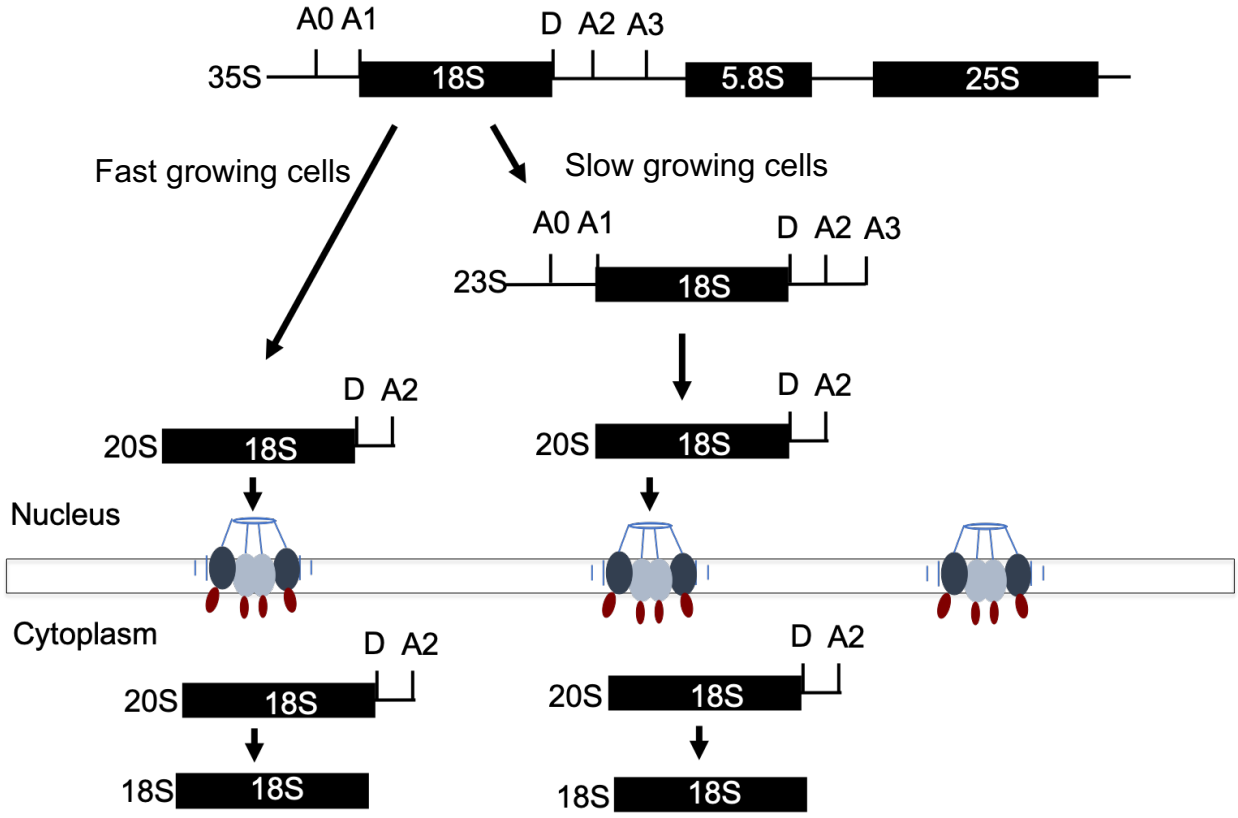


Figure 1-13. 18S rRNA processing. In fast-growing cells, co-transcriptional cleavage at A0, A1, and A2 sites generates the 20S pre-rRNA (left). In slow growing cells, rRNA processing occurs post transcriptionally after the synthesis of the 35S pre-rRNA, which undergoes cleavage at the A3 site to generate a 23S pre-rRNA. This is followed by cleavage of 23S pre-rRNA at A0, A1, and A2 sites to generate a 20S pre-RNA (right). In both instances, the 20S pre-rRNA is exported to the cytoplasm and is cleaved at the D site to generate mature 18S rRNAs.

40S subunit. If errors occur during processing, the inappropriately processed or assembled ribosomes are targeted by quality control systems (Cole *et al.*, 2009). This involves exosome mediated degradation, as loss of exosome function is known to result in the accumulation of 23S, 21S, and A2-C2 rRNAs, which are normally undetectable (Allmang *et al.*, 2000).

1.9.3 Maturation of 5.8S and 25S rRNAs

As with the 18S rRNA, the 25S and 5.8S rRNAs are produced by a set of nucleolytic cleavages from the 35S pre-rRNA (Figure 1-14) (Reeder *et al.*, 1999). Initial cleavage at the A2 and B0 sites produce a 27S-A2 pre-25S rRNA, which undergoes further cleavage at either the A3 or B1L site to produce 27S-A3 or 27S-B1L (Veldman *et al.*, 1980). The exonuclease Rat1p then acts to convert the 27S-A3 to 27S-B1S (Henry *et al.*, 1994). Both 27S-B1S and 27S-B1L undergo cleavage at the C2 site to produce 7S\L pre-5.8 rRNA and release the 26S pre-25S rRNA. These transcripts are then converted to 6S\L and 25S rRNAs by the activity of the exosome and Rat1p (Klinge and Woolford, 2019). Defects in the 3′ end processing of the 5.8 S rRNA was the original basis for characterizing the exosome as a 3′-5′ exonuclease (Mitchell *et al.*, 1996). *In vitro* and *in vivo* analyses demonstrate that 5.8 S rRNA processing also requires Mtr4p of the TRAMP complex (de la Cruz *et al.*, 1998).

1.10 Polyadenylation of noncoding RNAs

rRNAs synthesized by RNAPI are not polyadenylated, but previous work has shown that in wild-type cells a small portion of rRNAs have a poly(A) tail, which is significantly increased upon deletion of the exosome subunit Rrp6p (Kuai *et al.*, 2004). It is now appreciated that polyadenylation of rRNAs by Trf4p/Tr5p of the TRAMP complex mediates processing and degradation of rRNA pre-cursors and their by-products (e.g. 5′-ETS) generated by rRNA processing (Wery *et al.*, 2009; Weir *et al.*, 2010). In addition to RNAPI transcripts, exosome mutants also stabilize other polyadenylated ncRNA species, including RNAPIII tRNA and 5S

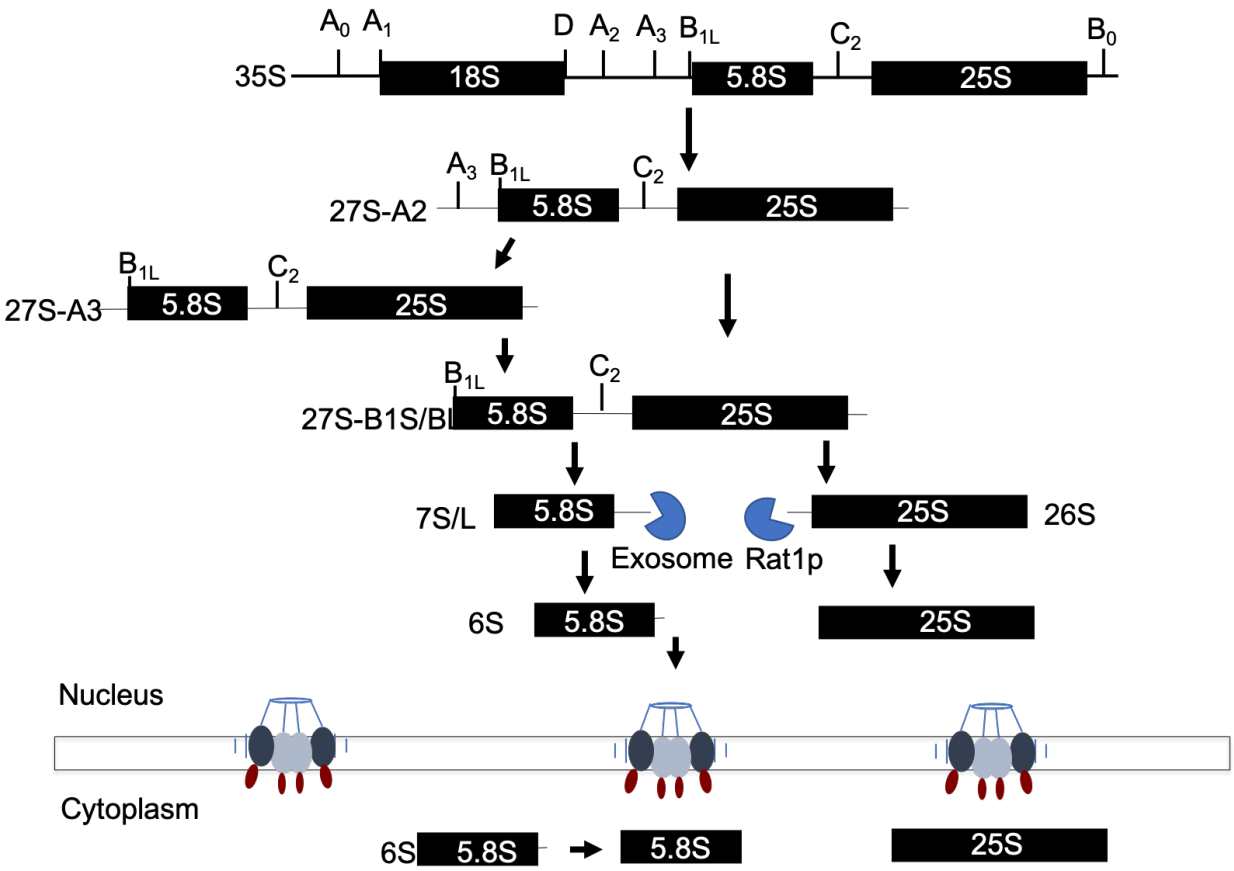


Figure 1-14. 25S rRNA processing. Co-transcriptional cleavage at the A2 site generates a 27SA2 pre-rRNA, which undergoes cleavage at A3 to generate a 27S-A3. 27S-A3 is trimmed by the exonuclease, Rat1p, to produce 27SB1S. On the other hand, direct cleavage at B1L generates 27SB1L. 27SB1S/L undergoes cleavage at the C2 site to generate 7S/L and 26S pre-rRNA. 7S/L undergoes trimming by the exosome to produce 6S, whereas 26S undergoes trimming by Rat1p to produce mature 25S rRNA. The 6S pre-rRNA is exported to the cytoplasm for further processing to produce mature a 5.8S rRNA.

rRNA transcripts and RNAPII snoRNA transcripts (Bertrand *et al.*, 1998; Ciganda and Williams, 2011). Although normally present at very low levels due to their transient nature, interruption of normal ncRNA processing steps required for removal of the poly(A) tail, or degradation of the poly(A)-tagged RNA can lead to the accumulation of large amounts of polyadenylated ncRNA in cells (Houseley and Tollervey, 2006; Wery *et al.*, 2009).

1.11 RNA processing and nuclear homeostasis

RNA processing in the nucleus is carried out through the dynamic interaction of RNAs with RBPs. As described above for all classes of RNAs, this involves a series of discrete events, with one event leading to the next through the RBPs activities and the associated cellular machinery that often engage each other. Various classes of RNAs also undergo polyadenylation either through the activity of canonical or non-canonical poly(A) polymerases. Generally, those poly(A) tails on rRNAs and snoRNAs are transient in nature and are removed as a part of their maturation process. In contrast, poly(A) tails on mRNAs are stable and core to the engagement of the mRNA by the gene expression apparatus. Bound to poly(A) tails are PABPs, that bind in a sequence specific manner (Brockmann *et al.*, 2012). Consequently, PABPs could indiscriminately bind both the poly(A) tails of mRNAs and ncRNAs, if present. These facts suggest that maintaining the proper distribution of poly(A) tails between mRNAs and ncRNAs, and the overall levels of nuclear poly(A)-RNA, would be central to PABP function. These concepts are supported by a recent study that showed that the nuclear accumulation of poly(A)-RNA in an mRNA export mutant led to Nab2p sequestration and rapid degradation of nascent transcripts, which could be reversed by overexpression of Nab2p (Tudek *et al.*, 2018b). Building off this observation, it is possible that any perturbation to RNA biogenesis or decay that results in altered patterns or amounts of poly(A)-RNA could have a similar impact on PABP availability. Moreover, in the absence of efficient clearance of accumulating poly(A)-RNAs, the titration of PABPs away from ongoing RNA processing events would be expected to cause further errors in processing and the generation of more substrates for decay by the RNA the surveillance machinery. This may reach a point where the exosome and other surveillance machineries are saturated with substrates, causing RNA surveillance and decay processes to fail. The positive

feedback loop established in this way would ultimately cause common terminal phenotypes via loss of nuclear RNA homeostasis, but in this model it is important to note that the initiating event could be different (e.g. export block or a change in ncRNA biogenesis).

1.12 Thesis Summary

In the work presented here, a screen using the budding yeast *S. cerevisiae* identified a set of genes, that when disrupted, caused poly(A)-RNA and mRNAs to accumulate in discrete locations within the nucleus (see Chapter III). The genes identified by this screen are known to function in diverse biological processes, including chromosome segregation, nucleocytoplasmic transport, RNA surveillance/decay, splicing, and ribosome biogenesis. It is not expected that all of these factors directly function in mRNA export; rather loss of some of these essential factors is likely to cause initial cellular defects that subsequently lead to poly(A)-RNA accumulation. For example, a number of factors involved in chromosome segregation were identified in the screen, which through loss of a chromosome carrying an essential mRNA export factor, could cause the observed phenotype. More interestingly, a group of genes that are involved in RNA surveillance/decay and ribosome biogenesis accumulated RNA and RBPs in the nucleolus when mutated. In chapter IV, for the ribosome biogenesis mutant *enp1-1*, the mechanism leading to the observed mRNA export block is shown to be mediated through a defect in the processing of nuclear polyadenylated rRNAs, which ultimately leads to a generalized failure in nuclear RNA homeostasis. A defect that is driven by sequestration of the nuclear poly(A) binding protein Nab2p by polyadenylated ncRNAs. Together this work identified poly(A)-RNA accumulation phenotypes that result from the disruption of essential genes in yeast, and the toxic effect of nuclear poly(A)-RNA accumulation on the cell. This included the identification of 15 genes with a novel poly(A)-RNA accumulation phenotype and a model of nuclear RNA homeostasis that involves maintaining a balance of RBPs, poly(A)-RNA, and RNA decay for a functional gene expression program.

Chapter II: *Experimental Methods*

2.1 Yeast strains, media, and plasmids

Yeast strains and plasmids used in this work are listed in Tables 2-1 and 2-2. Yeast mutant collections used in this work have been described previously (Ben-Aroya *et al.*, 2008; Breslow *et al.*, 2008; Li *et al.*, 2011). All yeast strains were grown in YPD (1% yeast extract, 2% bactopeptone, 2% glucose) at 25°C overnight unless otherwise stated. Strains containing autonomously replicating plasmids were grown in synthetic dropout medium (0.58 g Drop Out powder, 6.7 g Yeast Nitrogen Base without amino acids with ammonium sulfate, 20 g Agar, 100 ml 20% Dextrose, 100 ml 10X Amino Acid Drop out solution, and water to 1000 ml,) lacking the appropriate amino acid with 2% glucose.

Endogenous gene tagging with green florescent protein (GFP) was done for *NOP56*, *NAB2*, *PRP19*, and *HRP1* by amplifying each gene with the GFP::HIS3MX6 cassette plus ~300 bps of flanking sequence directly from the yeast GFP collection (Huh *et al.*, 2003). The resulting PCR product was transformed into yeast to integrate the GFP tag into the yeast genome (Gietz *et al.*, 1992). Deletion strains were made by homologous recombination with a selectable marker (*HIS3*) made by PCR (Longtine *et al.*, 1998). NDC1-GFP and GAL1-GFP-LDB19 strains were generated by transforming a PCR product with homology to the gene locus made from pKT148 (Sheff and Thorn, 2004) and pFA6a-His3MX6-PGAL1-GFP (Longtine *et al.*, 1998), respectively.

The NOP56-GFP plasmid (pBM461) was made by cloning NOP56-GFP with ~400 bp of promoter sequence from the yeast GFP collection (Huh *et al.*, 2003) into pRS313 (Sikorski and Hieter, 1989) using *XhoI* and *NotI* sites introduced by PCR. Tagging of the *ACT1* locus with a lacO array was performed in a wild-type strain as previously described using pSR13 and pAFS78 (Straight *et al.*, 1996; Rohner *et al.*, 2008), into which mutant alleles of select genes (i.e. *LDB19*, *SLI15,* and *SPC24*) were introduced by transformation with a PCR product. Rescue plasmids were either taken from the Yeast ORF Collection (i.e. *CEP3*, *ENP1*, *MPS1*, and *PRP2*) (Gelperin *et al.*, 2005), generated by cloning the gene (i.e. *ALR1*, *CLP1*, and *RRP43*) ± 500 bps of flanking sequence into pRS41N using *SacI* and *HindIII* sites introduced by PCR or obtained from the laboratory of Vivien Measday (University of British Columbia) (i.e. *CBF2* and *IPL1*) or Doug Koshland (UC Berkeley) (i.e. *SMC1*, *SMC3*, and *SMC4*). Overexpression plasmids for Pab1p-

GFP, Nab2p-GFP, and Yra1p-GFP were constructed by introducing the gene tagged at the C-terminus with GFP into a 2-μ plasmid (pBM005).

Table 2-1: Yeast Strains

Strain	Genotype	Reference
BMY 362	*MATa his3Δ1 leu2Δ met15Δ ura3Δ sli15-3::KANMX NOP56-GFP::HIS3MX6*	This study
BMY 662	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ ura10D::KANMX NOP56-GFP::HIS3MX6*	This study
BMY 664	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dis3-1::KANMX NOP56-GFP::HIS3MX6*	This study
BMY 667	*MATa his3Δ1 leu2Δ met15Δ ura3Δ smc4-1::KANMX NOP56-GFP::HIS3MX6*	This study
BMY 747	*MATa his3Δ1 leu2Δ met15Δ ura3Δ smc3-42::KANMX NOP56-GFP::HIS3MX6*	This study
BMY 640	*MATa his3Δ1 leu2Δ met15Δ ura3Δ mex67-5::KANMX ccw12Δ::HIS3MX6*	This study
BMY 710	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ ura10D::KANMX NDC1-GFP::HIS3MX6*	This study
BMY 703	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dbp5-1::KANMX NDC1-GFP::HIS3MX6*	This study
BMY 702	*MATa his3Δ1 leu2Δ met15Δ ura3Δ mex67-5::KANMX NDC1-GFP::HIS3MX6*	This study
BMY 704	*MATa his3Δ1 leu2Δ met15Δ ura3Δ prp2-1::KANMX NDC1-GFP::HIS3MX6*	This study
BMY 693	*MATa his3Δ1 leu2Δ met15Δ ura3Δ ura10D::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 658	*MATa his3Δ1 leu2Δ met15Δ ura3Δ rsp5-3::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 696	*MATa his3Δ1 leu2Δ met15Δ ura3Δ ldb19 Δ::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 695	*MATa his3Δ1 leu2Δ met15Δ ura3Δ spc24-10::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 694	*MATa his3Δ1 leu2Δ met15Δ ura3Δ sli15-1::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 745	*MATa his3Δ1 leu2Δ met15Δ ura3Δ nup133D::KANMX ACT1::256xLacO GFP-LacI::HIS3MX6*	This study
BMY 706	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ ura10D::KANMX NAB2-GFP::HIS3MX6*	This study
BMY 711	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dbp5-1::KANMX NAB2-GFP::HIS3MX6*	This study
BMY 713	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dis3-1::KANMX NAB2-GFP::HIS3MX6*	This study

BMY 712	*MATa his3Δ1 leu2Δ met15Δ ura3Δ rsp5-3::KANMX NAB2-GFP::HIS3MX6*	This study
BMY 709	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ ura10Δ::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 723	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dbp5-1::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 724	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dis3-1::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 748	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ rsp5-3::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 707	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ ura10D::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 719	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dbp5-1::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 722	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dis3-1::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 720	*MATa his3Δ1 leu2Δ met15Δ ura3Δ rsp5-3::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 718	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ rrp6D::KANMX NAB2-GFP::HIS3MX6*	This study
BMY 725	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ rrp6 Δ::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 721	*MATa, his3Δ1 leu2Δ met15Δ ura3Δ rrp6 Δ::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 734	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::KANMX NAB2-GFP::HIS3MX6*	This study
BMY 736	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::KANMX PRP19-GFP::HIS3MX6*	This study
BMY 735	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::KANMX HRP1-GFP::HIS3MX6*	This study
BMY 093	*MATa ura3-1 leu2-3 his3-11,15 trp1-1 ade2-1 rrp6 Δ::URA3*	This study
BMY 744	*MATa his3Δ1 leu2Δ met15Δ ura3Δ pGAL-GFP-LDB19::HIS3MX6*	This study
BMY 743	*MATa his3Δ1 leu2Δ met15Δ ura3Δ pGAL-GFP-LDB19::HIS3MX6 xpo1-1::KANMX*	This study
BMY 562	*MATa his3Δ1 leu2Δ met15Δ ura3Δ ura10Δ::URA3 RRP41-PrA::HIS3MX6*	This study
BMY 563	*MATa his3Δ1 leu2Δ met15Δ ura3Δ csl4-ph::URA3 RRP41-PrA::HIS3MX6*	This study
BMY 859	*MATa his3Δ1 leu2Δ met15Δ ura3Δ srm1-ts::KANMX RRP41-PrA::HIS3MX6*	This study
BMY 861	*MATa his3Δ1 leu2Δ met15Δ ura3Δ mtr4-1::KANMX RRP41-PrA::HIS3MX6*	This study
BMY 862	*MATa his3Δ1 leu2Δ met15Δ ura3Δ dis3-1::KANMX RRP41-PrA::HIS3MX6*	This study

BMY 863	*MATa his3Δ1 leu2Δ met15Δ ura3Δ mex67-5::KANMX RRP41-PrA::HIS3MX6*	This study
BMY 880	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::KANMX RRP41-PrA::HIS3MX6*	This study
BMY 1470	*MATa his3Δ1 leu2Δ met15Δ ura3Δ csl4-ph::URA3 TRF5 Δ::hygro*	This study
BMY1471	*MATa his3Δ1 leu2Δ met15Δ ura3Δ ura10D::URA3 TRF5 Δ::hygro*	This study
BMY1468	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::URA3 TRF5Δ::hygro*	This study
BMY 539	*MATa his3Δ1 leu2Δ met15Δ ura3Δ ura10 Δ::URA3*	This study
BMY 537	*MATa his3Δ1 leu2Δ met15Δ ura3Δ csl4-ph::URA3*	This study
BMY 547	*MATa his3Δ1 leu2Δ met15Δ ura3Δ csl4-ph::URA3 NAB2-PrA::HIS3MX6*	This study
BMY 789	*MATa his3Δ1 leu2Δ met15Δ ura3Δ enp1-1::KAMMX NAB2-PrA::HIS3MX6*	This study
BMY 1056	*MATa his3Δ1 leu2Δ met15Δ ura3Δ srm1::URA3 RRP41-GFP::HIS3MX6*	This study
BMY 544	*MATa his3Δ1 leu2Δ met15Δ ura3Δ CSL4::URA3 RRP41-PrA::HIS3MX6*	This study

Table 2-2: List of Plasmids

Name	Description	Reference
pBM461	pRS313-NOP56-GFP (HIS / CEN plasmid)	This study
pBM589	pFA6a-His3MX6-PGAL1-GFP	(Rohner *et al.*, 2008)
pBM590	pSR13-LacO-LEU2	(Rohner *et al.*, 2008)
pBM137	pRS313-TRF4-GFP (HIS / CEN plasmid)	This study
pBM550	pRS313-NRD1-GFP (HIS / CEN plasmid)	This study
pBM699	pRS423-NAB2-GFP (HIS / 2μ plasmid)	This study
pBM766	pRS423-PAB1-GFP (HIS / 2μ plasmid)	This study

2.2 Yeast transformation

Yeast cultures were grown overnight in YPD and diluted to an OD_{600} of 0.2 in fresh YPD, and then allowed to grow for another 4-6 hours or until the OD_{600} was 1.0. For each

transformation, 2 ml of culture was pelleted by centrifugation at 3300 relative centrifugal force (RCF), washed with sterile water, pelleted again, and then the following components were added in order: 240 µl of 50% PEG, 36 µl of 1.0 M lithium acetate, 10 µl of salmon sperm DNA (ssDNA) (10 mg/ml) and 64 µl of transformation DNA in water. Cells were mixed by vortexing for 2 minutes, incubated at 30°C for 30 minutes, and heat shocked for 15 minutes at 42°C. After cells were left to recover for two minutes at room temperature, cells were collected by centrifugation at 400 RCF for 3 minutes, re-suspended in 250 µl of sterile water, and spread onto selective medium to allow growth of transformed cells. Genotypes of select transformants were confirmed by PCR and sequencing when appropriate.

2.3 Yeast Growth Assay

Yeast cultures grown in selective medium at 25°C were diluted to a final OD_{600} of 0.1 in selective medium and 750 µl of each were added to individual wells of a 48-well plate. Plates were incubated in a Neo2 plate reader (BioTek) at 32°C with shaking for 1 minute prior to reading absorbance at OD_{600} every 10 minutes. Growth rates were calculated using the maximum slope of the resulting growth curves from 20 consecutive data points for three replicates of each strain.

2.4 RNA Polymerase I shut off assay

In order to stop rRNA synthesis, RNAPI activity was inhibited using the drug rapamycin. Both control and mutant strains were grown overnight at room temperature in YPD. Culture were diluted in the morning and were grown until an OD_{600} of 0.8. 2 ml of culture of each strain was transferred to fresh YPD. Rapamycin was added to one set of cultures to a final concentration of 100 ng/ml just before temperature shift and another set of cultures were treated with only dimethyl sulfoxide (DMSO) as a control. After 90 minutes of incubation at 37° C, cells were fixed in 5% formaldehyde (final concentration) and fluorescence *in situ* hybridization (FISH) was performed to check poly(A) localization.

2.5 Poly(A)-RNA FISH

Poly(A)-RNA was detected by FISH using a fluorescein isothiocyanate (FITC) labeled oligo-dT probe as previously described (Cole *et al.*, 2002). Briefly, yeast strains were grown in YPD at room temperature until OD_{600} = 0.8, transferred to 37°C for 90 minutes if necessary, and fixed by adding 37% formaldehyde (Sigma-Aldrich) directly to the culture to a 5% final concentration. Cells were fixed at the growth temperature (e.g. room temperature or 37°C) for 15 minutes, centrifuged at 3300 RCF for 1 minute and cell pellet was washed two times with cold buffer A (0.1 M potassium phosphate pH 6.5 and 0.5 mM $MgCl_2$). To remove the yeast cell wall, cells were then re-suspended with 250 µl of buffer B (0.1 M potassium phosphate pH 6.5, 0.5 mM $MgCl_2$, and 1.2 M Sorbitol) with beta-mercaptoethanol (Fisher) added to 0.4% final concentration and Zymolase at 0.25 mg/mL followed by incubation at 37°C for 35 minutes. In the meantime, eight well glass slides were coated by treating each well with 25 µl of 0.01% poly-L-lysine for 10 minutes and followed by drying the slide at room temperature. After spheroplasting, the cell pellet was collected by centrifugation at 400 RCF for 3 minutes and re-suspended in 200 µl of buffer B. Cells were then placed on poly-L-lysine coated glass slides and incubated at room temperature for 10 minutes. Un-adhered cells were removed by washing two times with buffer A, followed by permeabilization in ice-cold methanol for 6 minutes and ice-cold acetone for 30 seconds, after which slides were allowed to air dry. Prior to hybridization, the hybridization solution (3.8 ml diethylpyrocarbonate (DEPC) water, 5 ml 20X saline-sodium citrate (SSC), 1 ml 100X Denhardt's solution (1% ficoll, 1% polyvinylpyrrolidone, 1% bovine serum albumin), 200 µl 1% Tween 20) without probe was added to each well and incubated at 37°C for 2 hours and then replaced with hybridization solution containing 50 nucleotide long fluorescently labeled oligo dT probes and incubated at 37°C for at least 12 hours. The following day, cells were washed sequentially with 2X SSC, 1X SSC, 0.5X SSC for 30 minutes, and 2X phosphate buffered saline (PBS) for 10 minutes at room temperature, and after the final wash, slides were dipped into 100% ethanol for 10 seconds, air-dried, and mounting media (Vectorlabs) with 4' 6 diamidino-2-phenylidole (DAPI) was applied to each sample and a coverslip was affixed. Imaging was performed on a DeltaVision Elite microscope system equipped with a Front Illuminated sCMOS camera driven by Softworx 6 (GE Healthcare) using an Olympus 60X 1.42 N.A. oil objective. Image analysis and generation of cropped images for display were performed

in ImageJ using FIJI (Schindelin *et al.*, 2012). Identical imaging and display settings are used for all images presented in a figure panel.

2.6 Screen for mRNA export mutants by dT FISH

For the screening of the yeast *Ts* mutant collection alleles, yeast strains were obtained directly from each mutant collection (Ben-Aroya *et al.*, 2008; Breslow *et al.*, 2008; Li *et al.*, 2011), grown in YPD into log phase, and then shifted directly to 37°C for 3 hours. Imaging was performed on an Olympus IX81 microscope with 100X oil immersion objective (N.A.=1.4) controlled by Metamorph software (Molecular Devices Corp., Sunnyvale, CA) using identical exposure settings. During each day of screening, control*, mex67-5* and *dbp5-1* strains were included as controls to ensure consistency in the FISH procedure. Imaging data (>200 cells) was used to visually score nuclear accumulation of poly(A)-RNA with mutants showing evidence of accumulation being validated in triplicate. The specific *Ts* alleles in each strain found to accumulate poly(A)-RNA were subsequently verified by PCR to verify the presence of the mutation.

2.7 Gene specific FISH

Individual mRNA and rRNA transcripts were detected using a mixture of up to 48 fluorescently labeled oligonucleotide probes of 20 nucleotides in length (Biosearch Technologies Inc., Petaluma, CA) with a modified version of the Poly(A)-RNA FISH procedure (see above). All probe sequences used are listed in Table 2-3. Briefly, cells were fixed, spheroplasted, attached to a glass slide, and permeabilized, as in Poly(A)-RNA FISH. Cells were rehydrated with a hybridization solution (5X SSC, 5X Denhardt's solution, 0.1% Tween-20, 0.01 mg/mL ssDNA, 0.02 mg/mL *E. coli* tRNA, and 10 mM vanadyle ribonucleoside complex (VRC)) for 5 minutes and incubated with fresh hybridization buffer for at least one hour at 37°C. Hybridization buffer with 1 ng of fluorescein labeled locked nucleic acid (LNA) oligo-dT probe (Exiqon Inc., Woburn, MA), 20 ng of a Quasar 570 labeled gene specific probe, and/or 4 ng of Quasar 670

labeled probe against internal transcribed spacer sequence (*ITS*) of pre-rRNAs were added and incubated overnight (~14 hrs) at 37°C in a humidity chamber. The following day, cells were washed, prepared for imaging, and imaged as described in section 2.5.

Image analysis was performed in ImageJ using FIJI (Schindelin *et al.*, 2012) with each 3D single molecule FISH data set being reduced to a single maximum Z projection. mRNA FISH signals were identified as single points using noise tolerances set for each individual image and/or smFISH experiment to minimize the detection of false spots and compensate for sample variability. Masks of DAPI and *ITS* signals were used to quantify the number of mRNAs within compartments marked by these signals and to determine the number of cells within the image being analyzed. Average transcript number per cell was determined from each cell. In the control and most mutant strains, foci varied in intensity within a small range and were therefore counted as single transcripts. Bright nuclear foci were apparent in some mutants, but these were not counted in subsequent data analysis. This approach was taken for three reasons: (1) these foci were near transcription sites suggesting they are not mature transcripts or have not been released from the site of transcription, (2) the distribution of smFISH probes across the entire length of the mRNA does not allow the differentiation of multiple short transcripts from a full-length mRNA, and (3) these bright foci were generally rare.

Table 2-3: Gene Specific FISH probes

Transcripts		
GFA1	***ACT1***	***IMD2***
tctggatctttccactagat	aagcagcaacctcagaatcc	tcttgtagtctctaatggcg
catccactaaggtgtcgata	ccagaaccgttatcaataac	cttggtaggctcttggtaaa
catcgatagcaataccggtg	ggcaaaaccggctttacaca	aattcctgcactgacaaacc
aagcactcactttaccgatt	aagacagcacgaggagcgtc	tataagtcaacccacctctg
tgcttagtaatctcctcttt	tggtctaccgacgatagatg	gctaacttcagaggacgcaa
agtaacgtctctgttcggat	cgaccatgataccttggtgt	aatattcctggttagcttgg
tctagtatgcgcaataccac	taggagtctttttgacccat	cgtgtccattggagaggaaa
agatctttgagggtgacagt	ggattgagcttcatcaccaa	aaagtggccatttctgactc
cgaccacaaattggtcttct	cgtaaagtcaagatacctct	aaaccgataccacccaacag
agcaatacactcggtatcgg	aataccgtgttcaattgggt	tggggtacagttatggtgaa
taagtcatgcccattttgta	tttccatatcgtcccagttg	cccattttcatagttcttga
taacccgtatgaaccttcta	gtagaaggtatgatgccaga	caacggtcgtagttggagaa
ggcgataacctcattaggat	ttctggggcaactctcaatt	gcaaatccatacttttcctt
accaatcagtaaaggggacc	cagtcaaaagaacagggtgt	catctgtcgtgacagggaag
ccacgaagtcgacttttagt	gatttagggttcattggagc	accaactttgcatttctctt

gttttcttcgggaaattcca	ttgagtcatcttttctctgt	agttgtcctcaacgaattgt
aatggaatttccggttgacc	cgtagaaggctggaacgttg	agggtttttggtcatgacat
cccaagccaaatgatttgtt	gacaaaacggcttggatgga	taatgtgataccttgtgcgc
gggaaccagcttcaaattca	agttctaccggaagagtaca	ctacccttttgattttctt
ttggcggcaattggtagtaa	ccggaatccaaaacaatacc	taccctttcatcaacaacc
cctggattgagaatgtctca	ggaacgacgtgagtaacacc	tcagttcgggaaagcataga
ttggagatccatcttctgat	ggtagagagaaaccagcgta	cgctaatgggtagttctgat
gccgcatccgaagaaacaaa	atcgattctcaaaatggcgt	taacagttgcttggtgttgg
tgagccaaatcgtcatcttc	agtcagtcaaatctctaccg	atagtcccaatagaagcacc
catgatctgagctaactcca	ccacgttcactcaagatctt	tagtctttctttatcagcgt
gtagattctggttgctcata	agcagtggtggagaaagagt	acccagcttttaccaataat
accacccaatatcactttat	tgtcacggacaatttctctt	tttccttgggatgaatccaa
ctgacaactggtaaccatgc	gacgtaacatagtttttcct	ccacttgagcatgttcaatt
catgcgatcatgatcagtct	tttcttgttcgaagtccaag	agacctgggaaactctcttt
agcacgagtagccaaacatg	gaagattgagcagcggtttg	acaacgttaccagcgatgac
ccacactaactgggatatct	gttcgtaggatttttcaatt	caccggcagcaatcaaattg
ggcattttctgtccagaaag	agtgatgacttgaccatctg	ccataccaattctcaaaccg
acgcatacatcgtctctgaa	ctctgaatctttcgttacca	ccacaagccataacttcttg
aatttagagccagcatggta	ggatggaacaaagcttctgg	acacattgtagacggctgta
caattccgacagttaaggct	ggcagattccaaacccaaaa	ggaacaccgaattggttagc
atatgaacaccacagtgggt	agttgtaagtagtttggtca	gttttgaacaccaccatcag
caacaccaatttcaggacca	tcgacatcacacttcatgat	gaaccaagagccaaagcttt
tcgatacacggtcatctgac	accgtataattccttacgga	cataccacccatcataacag
aatacctgcttaatttggcc	taccaccggacataacgatg	atattcacctggtgattcgg
ttttattcttggttccagc	gcaatacctgggaacatggt	ccttcaatcttttaccatct
cctttaattcagtcgcacag	gatttccttttgcattcttt	tcttttgcatggcgtcaatg
ggtaacctctacccaataac	tggaagatggagccaaagcg	gaggtagatgcattaccttt
tcaactcacctgccaaaaca	aggagcaatgatcttgacct	gcgaccaaaacactgtctga
tagagagtctctggtaccaa	agacggagtactttctttct	atggatcctttgtcaacgac
ctttcttgcagtaacttgct	gaagccaagatagaaccacc	ccgatgtcttgacaggaatg
gtttgcaggtcgattgattt	tccacatttgttggaaggta	ctttacccctttgaacatta
caatccctttattaacagcc	tcgtcgtattcttgttttga	gcagaagcggttctgaattc
cgacggtaacagatttagcc	ggtgaacgatagatggacca	aattatgaacgccaccttct
Transcripts		
LEU1	***CCW12***	***ITS***
aaagagttcttggacccttg	gcggcgatagaagcgaca	agagtatcactcactaccaa
catgaaccaagtgtctgtcg	agaagcgacagcggcgac	aggccagcaatttcaagtta
tcgaaagcttgtggagaggt	tggtaacgttagcagcgg	cgcagagaaacctctctttg
caactcttctgacctttctg	tcttggctgacagtagca	cgtacttgcattatacctca
cgcgaatcagtttgcttgat	gatggtgaccaaagtggt	ccgcagttggtaaaacctaa
tggcatcactcataccgaaa	acagacgtggtcttcaca	ttccaatacgctcagtataa
cagtggtacctggtaaagtg	gctggggagacagtttca	cgctctcttcttatcgataa
accgtgagtagaggtatgag	ggtagcggtggaaaccaa	ttaagaacattgttcgccta
caaaggccagcgaaccaaag	taacgtcatcgacggtga	
attgtttgagtggccaagac	ccaggtggtgtattgagt	

atgtgataccaggtgatagc	cttcagtggtcaatgggc	
acacatggacatacgtgctt	gaagtaccgttctttggg	
ttaatcatacctgctcttgc	ggtaactggagcagcagt	
tgcctttggtgtattggaaa	gtggtgttctttggagct	
tcaaagttttccagtaggca	gtgagttggagcagcaga	
gtcgacagcttcaatgttga	gcaccagtgtaagaggtg	
ccccatgtaatagttggaat	tggcaaagccttagcagc	
ttggtaaggcatcttgagga	gccaacaaagcaccagca	
ctttccataccagatttctt	caacaaagcagcggcacc	
tccaaacccatataggctaa		
caatacggccattggtacaa		
gctgcagcacttcttaaatc		
gcttgtttttgaccaaacc		
ttgaacaaccagcttctctc		
ggcatccaaaatatcagggt		
ccttgacgaccttcgaaatt		
tatctacgaagtgacctgcg		
aacctttggactactttggt		
tttcatcttcggaagtgacc		
taagcagcactttcaagctc		
tcattggcaatatcttgggg		
ctaggtttagcaggtgtatc		
caaaaatggcttcatgccgg		
tctaatggtgcgctaatacc		
tgctttgggataatagcgtc		
aacccgttctcttaattgtc		
ctttacggaaacgccattca		
ttcttgatccttaccttgat		
cctccaaggttcaacattta		
tgcgtgttctctggaagaac		
caaagtcttttagagcccat		
tatcaccgtaagaaggtgca		
cccttattagcaatagggat		
tggtaaatcgacgcagagct		
ttaactaggcagtgctttct		
tctctcttctcaaagcttcg		
aattttgatccaccttccaa		
aatcctggtggactttatcg		

2.8 Imaging mRNA processing factors

Yeast strains carrying integrated GFP reporters were grown in selective media overnight at room temperature into log phase and then shifted directly to 37°C for 3 hours. Following the

temperature shift, cells were fixed for 15 minutes by mixing an equal volume of cells in growth media with a 4% paraformaldehyde solution. Following fixation, cells were processed for Poly(A)-RNA FISH as described in section 2.5 using 2 ng of a TYE 563 labeled LNA oligo-dT probe (Exiqon Inc., Woburn, MA) and 4 ng of Quasar 670 labeled *ITS* probe (Biosearch Technologies Inc., Petaluma, CA) by FISH (Cole *et al.*, 2002). Imaging was performed on a DeltaVision Elite microscope system equipped with a Front Illuminated sCMOS camera driven by Softworx 6 (GE Healthcare) using an Olympus 60x 1.42 N.A. oil objective or an Andor Dragonfly using a Leica DMi8 microscope and an Andor EMCCD camera driven by Fusion software using a 60X 1.4 N.A. oil objective. Image analysis was performed in ImageJ using FIJI (Schindelin *et al.*, 2012). All images are set to same minimum and maximum intensity, representative parts of the image were cropped and presented.

2.9 Northern blotting

To isolate RNA for northern blotting, log phase yeast cell cultures were shifted to 37°C for 90 minutes and then cells were rapidly frozen in liquid nitrogen and cryolysis was performed by solid phase milling in a planetary ball mill (Retsch) producing a fine cell grindate, as previously described (Oeffinger *et al.*, 2007). RNA extractions were performed using a hot phenol RNA extraction (Collart and Oliviero, 1993). Poly(A)-RNA was enriched from 770 µg of total RNA using PolyATtract mRNA Isolation System as per the supplier's specifications (Z5310, Promega). To isolate Nab2p-Protein A associated RNAs, affinity purifications were performed from 1 g of cell grindate with 0.2X RNasin (Oeffinger *et al.*, 2007). Isolated RNA pellets were re-suspended in formamide to be stored until use. For northern blotting analysis, 3 micrograms of total RNA, all isolated poly(A)-enriched RNA, and all Nab2p-Protein A associated RNA was resolved on a 1% agarose-formaldehyde gel in Tricine-Triethanolamine and transferred onto a nylon membrane in 10X SSC by capillarity force (Mansour and Pestov, 2013). The oligonucleotides probes used to detect the targeted RNA species were radioactively labelled using polynucleotide kinase and ATP (γ-^{32}P). All probe sequences used are listed in Table 2-4.

Table 2-4: List of probes used for Northern blotting

Name	Description	Sequence
004	5′ ITS downstream of D	5′-CGGTTTTAATTGTCCTA-3′
007	25S rRNA	5′-CTCCGCTTATTGATATGC-3′
008	18S rRNA	5′-CATGGCTTAATCTTTGAGAC-3′
017	5.8S rRNA	5′-GCGTTCTTCATCGATGC-3′
20	5.8S/*ITS2* boundary	5′-TGAGAAGGAAATGACGCT-3′
033	5′ ETS 278 nt downstream of A0	5′-CGCTGCTCACCAATGG-3′
041	5S rRNA	5′-CTACTCGGTCAGGCTC-3′
400	*ACT1* pre-mRNA	5′-CGATGGGTTCGTAAGCGTACTCCTACCGTGG-3′
403	*ACT1* mRNA	5′-TCTTGGTCTACCGACGATAGATGGGAAGACAGCA-3′
yU14	U14 snoRNA	5′-CGATGGGTTCGTAAGCGTACTCCTACCGTGG-3′
snR30	snR30 snoRNA	5′-GGAATATACTGCGGTAGGACGAAC-3′

2.10 ePAT assay

Yeast cultures were grown overnight at room temperature until an $O.D._{600}$ of 0.8 was reached, at which point cultures were shifted to 37°C for 90 minutes. Yeast cells were collected from 3 mL of culture by centrifugation (2 minutes at 2500 RCF) and total RNA was immediately isolated using phenol/chloroform/isoamyl alcohol (PCI) at a 25:24:1 (Collart and Oliviero, 1993). After isolating total RNA, DNA contamination was removed by DNaseI treatment followed by re-isolation of RNA using PCI. ePAT assays were performed as described previously (Janicke *et al.*, 2012) with each reaction using 1 µg of total RNA and the ePAT-anchor primer (BMO 711). Briefly, the mixture was incubated at 80°C for 5 min, cooled to room temperature, Superscript III enzyme buffer and dNTPs were added with DNA polymerase I Klenow (1 unit, source; New England Biolab) one hour at 25°C. Klenow was then inactivated by incubation of the mixture at 80°C for 10 minutes. After reducing the temperature to 55°C, Superscript III was added to the

reaction mix and incubated at 55°C for an hour to allow cDNA synthesis. Generated cDNAs were used as a template for gene specific PCR using primers targeting snR30 (BMO1012) and *APQ12* (BMO708) with the universal primer (BMO712). Probe sequences for each primer are listed in Table 2-5.

Table 2-5: List of primers used for ePAT

Probe No.	Sequence	Description
BMO711	GCGAGCTCCGCGGCCGCGTTTTTTTTTTTT	Anchor Primer
BMO712	GCGAGCTCCGCGGCCGCG	Universal primer
BMO708	GAAACGCCTCTGCTTACTCGG	*APQ12*
BMO1012	TCGGTCATCTTTGTTGTTCG	*SNR30*

2.11 Immunoprecipitation (IP)-Mass Spectrometry

Cells were grown at permissive temperature (25°C) until an OD_{600} of ~0. 8 before cultures were rapidly shifted to non-permissive temperature (37°C) by addition of pre-warmed media for 30 or 90 minutes prior to harvesting by centrifugation. Harvested cells were rapidly frozen in liquid nitrogen and cryolysis was performed by solid phase milling in a planetary ball mill (Retsch) producing a fine cell grindate (Oeffinger *et al.*, 2007). The grindate was stored at -80°C until processed for affinity purification. Nab2p-Protein A and Rrp41p-Protein A mRNP purifications, and a negative control (Protein A-tag expressed alone) were performed in RNP buffer (20 mM 4-(2-hydroxyethyl)-1-piperazinithane sulfonic acid (HEPES)-potassium hydroxide (HEPES-KOH) pH 7.4, 110 mM potassium acetate (KOAc), 0.5% Triton X-100, 0.1% Tween-20, 1:100 solution P, 1:5000 antifoam A) supplemented with 100 mM NaCl. Frozen cell grindate was rapidly thawed into RNP buffer and affinity purifications were carried out using magnetic IgG beads (Oeffinger et al, 2007). Following purification, on-bead digestion of the isolated complexes was performed in 20 mM Tris-HCl pH8.0 at 37°C using 1µg of trypsin per sample. The digestion was stopped by adding formic acid to a final concentration of 2% in a total

volume of 50 µL and tryptic digests were desalted using C18 ZipTips as per supplier recommendations (Millipore).

Liquid chromatography was performed using a PicoFrit fused silica capillary column (15 cm x 75 µm i.d; New Objective, Woburn, MA), self-packed with C-18 reverse-phase material (Jupiter 5 µm particles, 300 Å pore 503 size; Phenomenex, Torrance, CA) using a high-pressure packing cell. This column was installed on the Easy-nLC II system (Proxeon Biosystems, Odense, Denmark) and coupled to the LTQ Orbitrap Velos (ThermoFisher Scientific, Bremen, Germany) equipped with a Proxeon nanoelectrospray Flex ion source. The buffers used for chromatography were 0.2% formic acid (buffer A) and 100% acetonitrile/0.2% formic acid (buffer B). Peptides were loaded on-column at a flowrate of 600 nL/min and eluted with a 3 slope gradient at a flowrate of 250 nL/min. Solvent B was first increased from 2 to 25% over 20 minutes, then from 25 to 45% over 40 minutes, and finally from 45 to 80% B over 10 min. LC-MS/MS data was acquired using a data-dependant top method combined with a dynamic exclusion window of 22 seconds. The mass resolution for MS was set to 60,000 (at m/z 400) and used to trigger the sixteen additional MS/MS events acquired in parallel in the linear ion trap for the top eleven most intense ions. Mass over charge ratio range was from 370 to 1800 for MS scanning with a target value of 1,000,000 charges and from ~1/3 of parent m/z ratio to 2000 for MS/MS scanning with a target value of 10,000 charges. The data dependent scan events used a maximum ion fill time of 100 milliseconds (ms) and 1 microscan. Nanospray and S-lens voltages were set to 1.3–5171.7 kV and 50 V, respectively. Capillary temperature was set to 250ºC. MS/MS conditions used were normalized collision energy of 35 V; activation q of 0.25; and an activation time of 10 ms.

The peak list files were generated with Proteome Discoverer (version 2.1) using the following parameters: minimum mass set to 500 Da, maximum mass set to 6000 Da, no grouping of MS/MS spectra, precursor charge set to auto, and minimum number of fragment ions set to 5. Protein database searching was performed with Mascot 2.5 (Matrix Science) against the NCBI *S. cerevisiae* protein database (20160802). The mass tolerances for precursor and fragment ions were set to 10 ppm and 0.6 Da, respectively. Trypsin was used as the enzyme allowing for up to 1 missed cleavage. Cysteine carbamidomethylation was specified as a fixed modification, and methionine oxidation as variable modifications. Data analysis was performed using Scaffold (version 4.8.4).

Mass spectrometry results were analysed as per Scott and Trahan et al, 2017 (Scott *et al.*, 2017). Briefly, preys having an average of less than 2 unique peptide count from the Nab2p- or Rrp41p-Protein A pullout conditions were first filtered out. The average of spectral counts from all biological replicates, termed as average total spectral counts (averageTSC), was calculated. The highest total spectrum counts from all negative controls (hbgTSC from Protein A alone) were then subtracted from the average total spectrum of each prey (averageTSC –hbgTSC) under their respective conditions. The resulting average prey total spectrum count containing less than two total spectrum counts were again filtered out. To take into account the variable sizes of all proteins identified to normalize the peptide counts, all proteins were *in silico* digested with trypsin using MS-digest (http://prospector.ucsf.edu) using zero missed cleavages and no variable modifications to determine how many theoretical peptides each protein would generate upon trypsin digestion (in silico peptides). Average total spectrum counts of each protein were then divided by its theoretical number of peptides (averageTSC/*in silico* peptides) before normalizing the number of prey peptides to the peptide values for Nab2p-Protein A or Rrp41p-Protein A.

2.12 Representation of protein-protein interaction data

To aid in data visualization of the proteins identified in pulldowns with Rrp41p-Protein A by mass spectrometry, the individual proteins were manually clustered around associated GO terms (geneontology.org) using Cytoscape (Shannon *et al.*, 2003). This is displayed as nodes representing biological process (rectangles) connected to nodes representing proteins (circular) by an edge. In this representation, edge lengths or distances from another node are not meant to be informative, only the connection to the GO process is to be considered. For proteins identified with Nab2p-Protein A, heatmaps were generated from normalized averageTSC counts (see section 2.10) and were manually clustered according to their associated GO terms.

2.13 RNA- sequencing (RNA-seq)

Harvested cells were rapidly frozen in liquid nitrogen and cryolysis was performed by

solid phase milling in a planetary ball mill (Retsch) producing a fine cell grindate (Oeffinger *et al.*, 2007). RNA extractions were performed using 5 mg of cell grindate that was thawed into 500 µL of Trizol. All kits were used as per supplier specifications. The RNA extracts (74104, Qiagen) were either Poly(A)-RNA enriched via oligo dT (E7490, NEB) or ribo-depleted (MRZY1324, Epicentre) and cDNA libraries were prepared using the Kapa stranded RNA-seq library preparation kit (Kapa Biosystems). Paired end 50 (PE50) RNA-sequencing was performed using a Hiseq 2500.

2.14 RNA-seq analysis

Below is a description of the RNA-seq analysis pipeline with the associated commands that are needed to recapitulate the data analysis. Code related to this analysis is also available at https://github.com/montpetitlab/Paul_et_al_2019.

To start, Pair-End RNA-seq data were aligned to the reference genome assembly R64-1-1 (GCA_000146045.2) using HISAT2 (version 2.0.4) to generate outputs in a SAM format (Kim *et al.*, 2015). The resulting outputs from HISAT2 were piped to samtools and converted to BAM files.

```
"hisat2 -p 4 --fr -x path to genome file --known-splicesite-
infile -1 R1.fastq -2 R2.fastq | samtools view -Sb >
output_file.bam"
```

BAM files were then sorted and indexed using samtools (version 1.6) for visualization by IGV software (Robinson *et al.*, 2011) and to generate Figure 4-12.

```
"samtools sort input_file.bam -o output_file_sorted.bam"
"samtools index input_file.bam"
```

A count matrix was generated for all the samples by using the featureCounts function of R package Rsubreads (version 1.28.1) and a custom annotation file (Liao *et al.*, 2014). The custom annotation file was generated by combining annotation files from the ensemble assembly

with annotations for CUTs, SUTs, XUTs and NUTs (Xu *et al.*, 2009; Van Dijk *et al.*, 2011; Schulz *et al.*, 2013; Vera and Dowell, 2016). The R scripts used to generate the count matrix from bam files are given in Appendix B.

Differential expression was performed using the R package DESeq2 (Version 1.18.1) by applying the DESeq function (Love *et al.*, 2014). The R scripts for differential expression are provided in Appendix B. For these analyses, three biological replicates from *csl4-ph* and two from *rrp6Δ, dis3-1*, *enp1-1, srm1-ts*, and control strains were used. Replicates are used to measure dispersion of data in every sample, to calculate log2FC in the mutant strains compared to the control strain and perform statistical tests to generate an associated p-value. The Log2FC data generated from DESeq2 was filtered using a significance cut off of $p < 0.01$ to generate a list of differentially expressed genes for further analysis. In order to determine the similarity of log2FC between mutants, Pearson correlation coefficients were calculated using the cor.test() function of R which outputs a correlation co-efficient and associated p-value. P-values < 0.01 were again considered significant.

Metagene plots and heatmaps for snoRNA data panels were generated by using deepTools (Proudfoot *et al.*, 2002). The pipeline used was as follows: sorted BAM files were converted to bigwig files using the bamtobw command and the resulting bigwig files were used as input for the computeMatrix command. ComputeMatrix takes an annotation file for snoRNA and converts each snRNA gene annotation to an equal number of bins and calculates the number of reads in each bin to generate a matrix file. Subsequently, matrix files were used as inputs in to plotProfile and heatmap commands were used to generate Figure 4-12 panel B and C. Commands used are given below:

```
"bamCoverage -bs 1 -b input_sorted_indexed_bamfile.bam -o output.bw"
"computeMatrix scale-regions -R annotation.bed -S output.bw -b 200 -a 200 --skipZeros -o matrix.mat.gz"
"plotHeatmap -m matrix.mat.gz -out MetagenePlot.png"
"plotProfile -m matrix.mat.gz -out ExampleHeatmap.png"
```

Chapter III: *Altered RNA processing and export leads to retention of mRNAs near transcription sites, nuclear pore complexes, or within the nucleolus*

---

* A version of this chapter has been published as *Paul, B., & Montpetit, B. (2016). Altered RNA processing and export lead to retention of mRNAs near transcription sites and nuclear pore complexes or within the nucleolus. Molecular Biology of the Cell, 27(17), 2742–2756***.**

3.1 Introduction

As described in Chapter I, nascent transcripts are processed within the nucleus, which can include folding, cleavage, modification, nuclear export, or decay. Processing is driven by specific RNA–protein and RNA-RNA interactions that occur in the context of a RNP particle; the protein composition of the RNP is largely responsible for the processing path that is followed (Mitchell and Parker, 2014; Oeffinger and Montpetit, 2015; Singh *et al.*, 2015). Inevitably, errors occur during nuclear RNA processing, which can result in aberrant transcripts being targeted for nuclear degradation via quality control mechanisms (Houseley and Tollervey, 2009; Müller-McNicoll and Neugebauer, 2013; Eberle and Visa, 2014; Porrua and Libri, 2015). A key player in this process is the exosome, which functions in the nucleus and cytoplasm as a nuclease to facilitate and survey RNA biogenesis from all three nuclear RNA polymerases (Chlebowski *et al.*, 2013; Porrua and Libri, 2013; Schneider and Tollervey, 2013).

Disruptions to RNA biogenesis, export, and surveillance results in the accumulation of aberrant RNP complexes and RNA processing by-products in the nucleus of the affected cell (Amberg *et al.*, 1992; Kadowaki *et al.*, 1992, 1994b; Doye *et al.*, 1994; Fabre *et al.*, 1994; Gorsch *et al.*, 1995; Segref *et al.*, 1997). For example, when exosome-dependent RNA processing and surveillance is perturbed the biogenesis of snoRNA, rRNA, tRNA, mRNA, and other noncoding transcripts are altered (van Hoof *et al.*, 2000a; Kuai *et al.*, 2004; Vanácová *et al.*, 2005; Wyers *et al.*, 2005; LaCava *et al.*, 2005; David *et al.*, 2006; Davis and Ares, 2006; Houalla *et al.*, 2006; Carneiro *et al.*, 2007; Gudipati *et al.*, 2012; Schneider *et al.*, 2012; Castelnuovo *et al.*, 2013). Disruptions to RNA biogenesis and export can be observed as the accumulation of poly(A)-RNA species within the nucleus of the affected cell (Cole *et al.*, 2002), which has been used to identify many mutants involved in RNA biogenesis, including screening of the ~5000 non-essential genes in *S. cerevisiae* for mRNA export defects (Hieronymus *et al.*, 2004). The recent construction of mutant libraries that span essential genes (Ben-Aroya *et al.*, 2008; Breslow *et al.*, 2008; Li *et al.*, 2011), provides an opportunity to conduct comprehensive screens of essential genes for mRNA processing and export defects, as recently performed for tRNAs (Wu *et al.*, 2015).

Importantly, the screening of both essential (this work) and non-essential (Hieronymus *et al.*, 2004; Thomsen *et al.*, 2008) genes within *S. cerevisiae* for mRNA processing defects is expected to provide a component list that is necessary for building improved models of mRNA

biogenesis and export. Towards this goal, a screen of essential gene mutants for nuclear poly(A)-RNA accumulation and the characterization of these mutants using single molecule FISH (smFISH) directed against specific mRNAs was performed. This resulted in the identification of 15 genes that were not previously linked and/or demonstrated to alter RNA processing and mRNA export. In addition, disruption of multiple nuclear processes was found to cause distinct phenotypes that included the accumulation of mRNAs near transcription sites, the nuclear periphery and NPCs, or within the nucleolus. These data suggest that alterations to RNA processing and overall nuclear homeostasis cause RNAs to stall or be retained at similar restriction points. This may reflect common failures in mRNA biogenesis and export, as well as, active mechanisms to protect the cell during cellular stress and dysfunction.

3.2. Results

3.2.1 Identification of mutants that accumulate nuclear poly(A)-RNA

To identify genes involved in mRNA biogenesis and export, two temperature sensitive (*Ts*) *S. cerevisiae* mutant collections (Ben-Aroya *et al.*, 2008; Li *et al.*, 2011) were screened for the accumulation of poly(A)-RNA in the nucleus using an oligo-dT FISH assay. Together these two *Ts* collections cover ~68% (785/1156) of essential genes. When a *Ts* allele was not available, a Decreased Abundance by mRNA Perturbation (DAmP) allele was used that harbors a disrupted 3′ UTR that often leads to reduced gene expression (Breslow *et al.*, 2008), which increased coverage of essential genes to ~91% (1047/1156) in our screen. For consistency throughout the screening process, regardless of the type of mutant used, all strains were grown into log phase at 25°C, shifted to 37°C for 3 hours, and fixed. A three-hour temperature shift was used to balance the time needed to induce the *Ts* mutant phenotype(s) while minimizing induction of secondary phenotypes caused by the loss of essential cellular activities. In the case of the DAmP alleles, it was reasoned that the temperature shift might act as a stress and exacerbate mutant phenotypes, although DAmP alleles are not necessarily *Ts* mutants. Following fixation, *in situ* hybridization assays were performed using a fluorescently labeled oligo-dT probe to detect poly(A)-RNA. By comparing the distribution of oligo-dT to the DAPI signal, 29 of 1047 mutants were identified to

accumulate poly(A)-RNA in the nucleus (Figure 3-1 and Table 3-1) Of the genes identified, approximately half (14/29) have previously been reported to display nuclear accumulation of poly(A)-RNA when disrupted (Table 3-1).

To verify that the poly(A)-RNA accumulation phenotype was linked to the purported mutant being screened, all strains were verified by PCR, and in the case of the 15 newly reported genes, the mutation was rescued by introducing a wild-type allele and/or recapitulated by moving the mutation to a different strain background (Table 3-1). Genes previously reported to accumulate poly(A)-RNA in the nucleus when disrupted were identified by our screen (Table 6-1, Appendix A), which may be due to the specific allele present in the mutant collection, the length of the temperature shift, or the requirement of a poly(A) tail for detection in the initial *in situ* screen. Within the set of mutants identified the distribution of poly(A)-RNA within the nucleus was distinct and included bright foci (*brl1-3231*), a diffuse nuclear signal (*rsp5-3*), a diffuse nuclear signal with one or more foci (*dbp5-1*), or poly(A)-RNA being adjacent to the DAPI stained DNA mass (*dis3-1*) (Figures 3-1 and 3-2A). The pattern of poly(A)-RNA accumulation was similar for genes with related biological functions.

3.2.2 Nucleolar disruption is linked to poly(A)-RNA accumulation

Studies in *Schizosaccharomyces pombe* (*S. pombe*) have shown that mutants affecting chromosome biology are associated with the accumulation of nuclear poly(A)-RNA and nucleolar disruption (Kalam Azad *et al.*, 2003; Ideue *et al.*, 2004). Of the genes identified in our screen, ~1/3 have functions that include kinetochore-microtubule attachment, chromosome organization, and cell cycle checkpoint control (Table 3-1). Given these facts and the reported defects in rDNA segregation within mutants identified here (e.g. *IPL1* and *SMC* genes) (Freeman, 2000; D'Amours *et al.*, 2004; Sullivan *et al.*, 2004; Machín *et al.*, 2005), we assayed nucleolar status in all 29 mutants based on the presence and localization of two nucleolar markers, an rRNA processing intermediate (via *ITS*) and a nucleolar protein associated with rRNA processing (Nop56p-GFP) (Gautier *et al.*, 1997; Woolford and Baserga, 2013). By monitoring each nucleolar marker and poly(A)-RNA in the same cell, we found that 21 of the 29 mutants displayed alterations in *ITS* and/or Nop56p-GFP

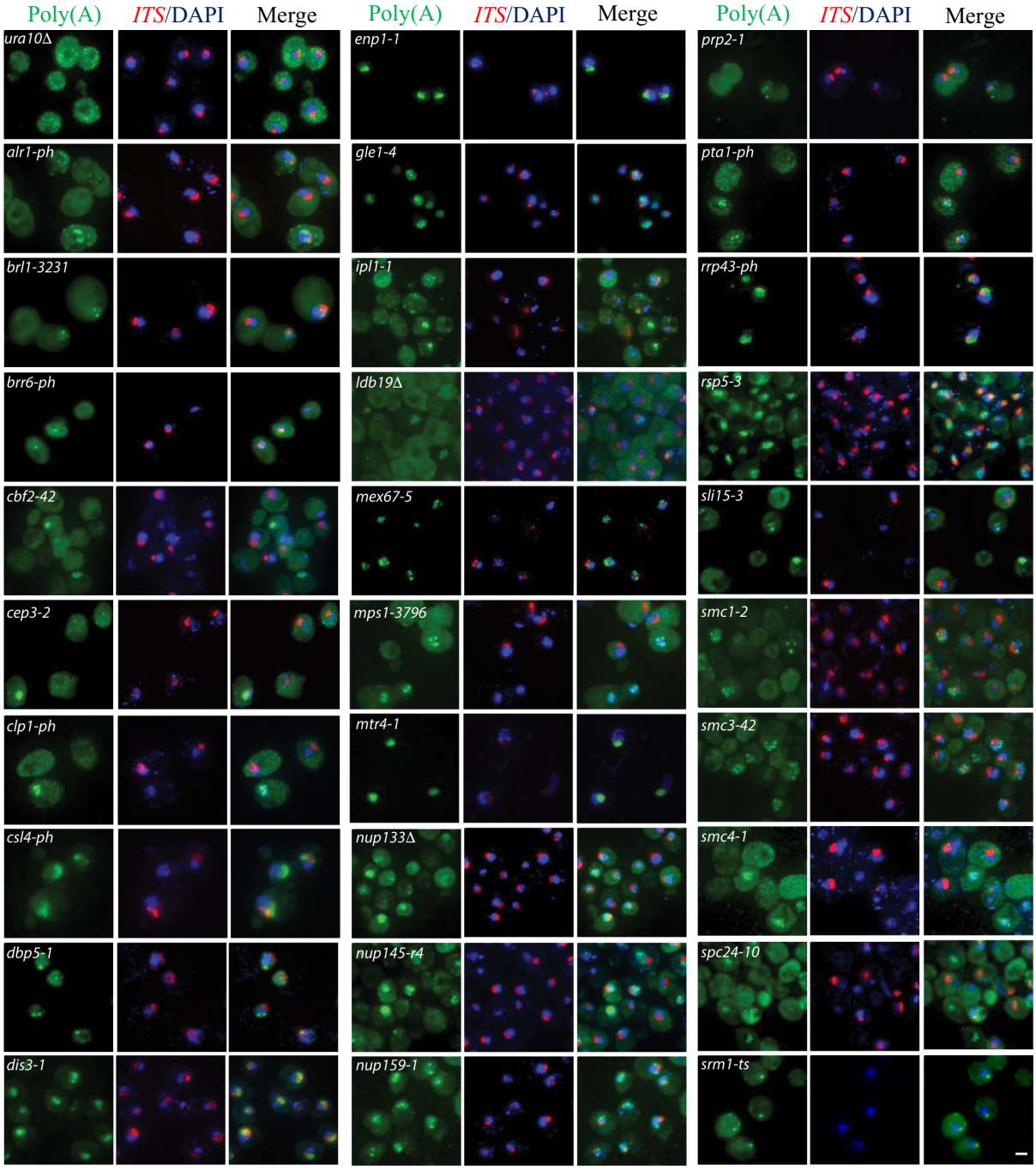


Figure 3-1. Nuclear poly(A)-RNA accumulation in *Ts* mutants. Representative images showing poly(A)-RNA (green) localization in control (*ura10Δ*) and the 29 mutants identified to have nuclear accumulation of poly(A)-RNA compared to *ITS* (red) and DAPI (blue) following 3 hours at 37°C. Scale bar = 1 μm.

Table 1: Description and phenotypes associated with mutants that display poly(A)-RNA accumulation.

Gene	Biological Process	Previous report / verification[1]	Distribution of nuclear poly(A)-RNA	Cells with nuclear poly(A)-RNA accumulation ± SD	*GFA1* mRNAs per cell ± SD	Nuclear *GFA1* mRNAs[2]	*ITS* and Nop56p-GFP
URA10	Pyrimidine biosynthesis	-	diffuse	<1%	5.1 ± 0.7	0.9 ± 0.9 (1)	-
ALR1	Plasma membrane Mg^{2+} transporter	rescued	foci	36 ± 2%	11.1 ± 1.7	1.8 ± 1.6 * (1)	-
BRL1	NPC and NE biogenesis	(Saitoh *et al.*, 2005)	diffuse signal / foci	34 ± 11%	6.0 ± 0.6	1.3 ± 1.1* (1)	-
BRR6	NPC and NE biogenesis	(de Bruyn Kops and Guthrie, 2001)	diffuse signal / foci	49 ± 4%	4.5 ± 1.1	1.1 ± 1.0* (1)	-
CBF2	Chromosome segregation	rescued	diffuse signal / foci	27 ± 11%	5.2 ± 1.2	0.9 ± 0.9 (1)	both absent in cells with nuclear poly(A)-RNA
CEP3	Chromosome segregation	rescued	diffuse signal / foci	16 ± 4%	6.7 ± 0.9	1.0 ± 1.3 (1)	both absent in cells with nuclear poly(A)-RNA
CLP1	Cleavage and polyadenylation of RNA	rescued	foci	32 ± 10%	9.6 ± 1.3	1.6 ± 1.3* (1)	Nop56p-GFP foci
CSL4	RNA processing and degradation	remade	diffuse signal next to DAPI	~100%	6.0 ± 1.9	2.3 ± 1.9* (2)	enlarged *ITS* and Nop56p-GFP area
DBP5	RNA export	(Snay-Hodge *et al.*, 1998)	diffuse signal / foci	~100%	4.2 ± 0.3	2.2 ± 1.5* (2)	*ITS* and Nop56p-GFP foci
DIS3	RNA processing/ degradation	(Kadowaki *et al.*, 1994a)	diffuse signal next to DAPI	~100%	5.5 ± 1.0	2.1 ± 1.5* (2)	enlarged *ITS* and Nop56p-GFP area
ENP1	RNA processing and ribosomal subunit synthesis	rescued	diffuse signal next to DAPI	~100%	3.4 ± 0.5	2.0 ± 1.5* (2)	*ITS* decreased or absent
GLE1	RNA export	(Murphy and Wente, 1996; Del Priore *et al.*, 1996)	diffuse signal / foci	~100%	2.2 ± 0.3	2.1 ± 1.3* (2)	*ITS* and Nop56p-GFP foci
IPL1	Chromosome segregation / cell cycle	(Cole *et al.*, 2002)	diffuse signal / foci	30 ± 12%	3.4 ± 0.5	0.8 ± 0.9 (1)	both absent in cells with nuclear poly(A)-RNA
LDB19	Ubiquitin-dependent endocytosis	remade	diffuse signal	17 ± 3%	6.2 ± 0.5	0.9 ± 1.0 (1	-
MEX67	RNA export	(Segref *et al.*, 1997)	diffuse signal / foci	~100%	2.4 ± 0.6	2.1 ± 1.2* (2)	*ITS* and Nop56p-GFP foci
MPS1	Spindle pole body / cell cycle	rescued	diffuse signal / foci	28 ± 4%	6.5 ± 1.2	1.1 ± 1.1 (1)	both absent in cells with nuclear poly(A)-RNA

MTR4	RNA processing and surveillance	(Kadowaki *et al.*, 1994a)	diffuse signal next to DAPI	~100%	4.9 ± 0.7	2.3 ± 1.5* (2)	*ITS* decreased or absent
NUP133	Nucleocytoplasmic transport	(Doye *et al.*, 1994)	diffuse signal / foci	59 ± 5%	5.7 ± 0.8	1.3 ± 1.0* (1)	-
NUP145	Nucleocytoplasmic transport	(Wente and Blobel, 1994; Fabre et al., 1994)	diffuse signal / foci	43 ± 2%	4.6 ± 0.6	1.2 ± 1.0* (1)	-
NUP159	Nucleocytoplasmic transport	(Gorsch *et al.*, 1995)	diffuse signal / foci	~100%	1.7 ± 0.2	1.6 ± 1.0* (1)	*ITS* and Nop56p-GFP foci
PRP2	pre-mRNA splicing	rescued	diffuse signal	34 ± 12%	5.9 ± 0.5	1.1 ± 1.1 (1)	-
PTA1	Cleavage and polyadenylation of RNA	(Hammell *et al.*, 2002)	diffuse signal / foci	35 ± 8%	6.1 ± 0.9	1.2 ± 1.0* (1)	-
RRP43	RNA processing and degradation	rescued	diffuse signal next to DAPI	~100%	5.3 ± 0.3	1.8 ± 1.7* (1)	enlarged *ITS* and Nop56-GFP area
RSP5	E3 ubiquitin ligase; multiple processes	(Neumann *et al.*, 2003; Rodriguez *et al.*, 2003)	diffuse signal	~100%	5.7 ± 2.0	1.2 ± 0.9* (1)	-
SLI15	Chromosome segregation/cell cycle	remade	diffuse signal / foci	30 ± 4%	4.9 ± 1.1	1.0 ± 1.2 (1)	both absent in cells with nuclear poly(A)-RNA
SMC1	Chromosome segregation	rescued	diffuse signal / foci	14 ± 2%	4.3 ± 0.5	1.0 ± 1.2 (1)	-
SMC3	Chromosome segregation	rescued	diffuse signal / foci	51 ± 13%	5.7 ± 0.3	1.3 ± 1.2* (1)	-
SMC4	Chromosome organization	rescued	diffuse signal / foci	29 ± 7%	9.0 ± 2.5	1.1 ± 1.3 (1)	both absent in cells with nuclear poly(A)-RNA
SPC24	Chromosome segregation	remade	diffuse signal / foci	33 ± 6%	4.3 ± 1.1	0.9 ± 1.1 (1)	both absent in cells with nuclear poly(A)-RNA
SRM1/ PRP20	Nucleocytoplasmic transport	(Kadowaki *et al.*, 1994a; Amberg *et al.*, 1993)	diffuse signal next to DAPI	~100%	2.1 ± 0.2	0.6 ± 0.7* (0)	*ITS* decreased or absent

[1] To verify mutants that had not previously been reported to accumulate poly(A)-RNA, the phenotype was rescued using a wild-type allele or recapitulated by making the mutant in a different strain background as indicated.
[2] Average number of nuclear mRNAs ± SD with the median for the data shown in brackets.
***** Distribution of nuclear mRNAs tested using the Wilcoxon-rank sum test and found to be significant at $p<0.001$.

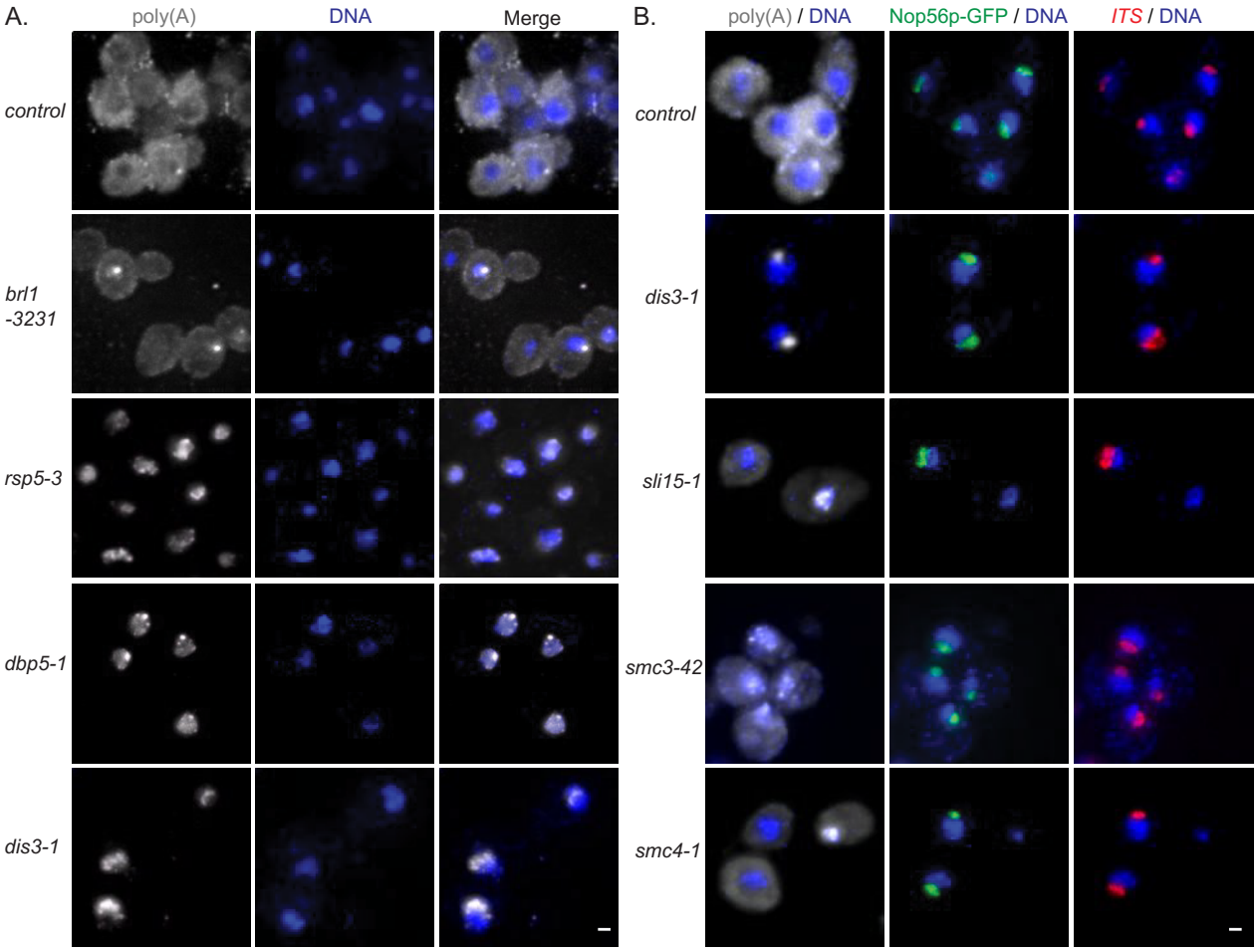


Figure 3-2. Poly(A)-RNA localization patterns and nucleolar status. (A) Representative images showing poly(A)-RNA (grey) localization in control (*ura10Δ*) and select mutant strains compared to DAPI (blue) following 3 hours at 37°C. (B) Representative images showing poly(A)-RNA (grey) localization in control (*ura10Δ*) and select mutant strains compared to DAPI (blue), Nop56p-GFP (green) and *ITS* (red) following 3 hours at 37°C. Scale bars = 1 μm.

localization and abundance when poly(A)-RNA accumulated (Table 3-1), which included fragmented nucleoli in mRNA export mutants as previously reported (Kadowaki *et al.*, 1994c; Dockendorff *et al.*, 1997; Segref *et al.*, 1997; Thomsen *et al.*, 2008). Nucleolar disruption was prominent in 7 of the 9 mutants that affect chromosome biology, with these mutants often lacking nucleolar *ITS* and Nop56p-GFP in cells with poly(A)-RNA accumulation (e.g. *sli15-1* and *smc4-1* vs *dis3-1,* Figure 3-2B). While *SMC1* and *SMC3* mutants did not show obvious nucleolar defects, recent reports provide a direct role for cohesins (e.g. *SMC3*) in nucleolar function (Bose *et al.*, 2012; Harris *et al.*, 2014). Together, these findings support a link between poly(A)-RNA accumulation and alterations to the nucleolus, which may often be induced by errors in chromosome segregation.

3.2.3 Identification of mRNA biogenesis and export mutants

Mutations within RNA processing and surveillance pathways have been shown to accumulate poly(A)-RNA species that included rRNA, mRNA, snRNA, snoRNA, and tRNA (van Hoof *et al.*, 2000a; Kuai *et al.*, 2004; LaCava *et al.*, 2005; Vanácová *et al.*, 2005; Wyers *et al.*, 2005; Carneiro *et al.*, 2007; Rougemaille *et al.*, 2007; Gudipati *et al.*, 2012; Schneider *et al.*, 2012; Castelnuovo *et al.*, 2013). Consequently, the mutants identified here using an oligo-d(T) based *in situ* hybridization approach may accumulate poly(A)-RNA due to disruptions in RNA biogenesis that are independent of mRNA. To identify those mutants that alter mRNA processing and export, single molecule FISH (smFISH) assays were performed using probes against *GFA1, ACT1,* or *CCW12* transcripts. The mRNAs were selected based on relative expression levels (*GFA1* = low and *ACT1*/*CCW12* = high) and the presence of an intron in *ACT1* (Ng and Abelson, 1980), which may lead to this mRNA being affected differently than non-spliced mRNAs (i.e. *CCW12* and *GFA1*). To ensure that a block in mRNA export could be observed, we employed a *mex67-5* strain that when shifted to the non-permissive temperature, robustly accumulated poly(A)-RNA in the nucleus (Segref *et al.*, 1997). smFISH assays using the gene specific probes displayed an obvious block in mRNA export at the non-permissive temperature in *mex67-5* and, importantly, a *mex67-5* / *ccw12Δ* strain showed no detectable signal with the *CCW12* probe set (Figure 3-3A). This established that these mRNA probes can be used to detect export defects, and

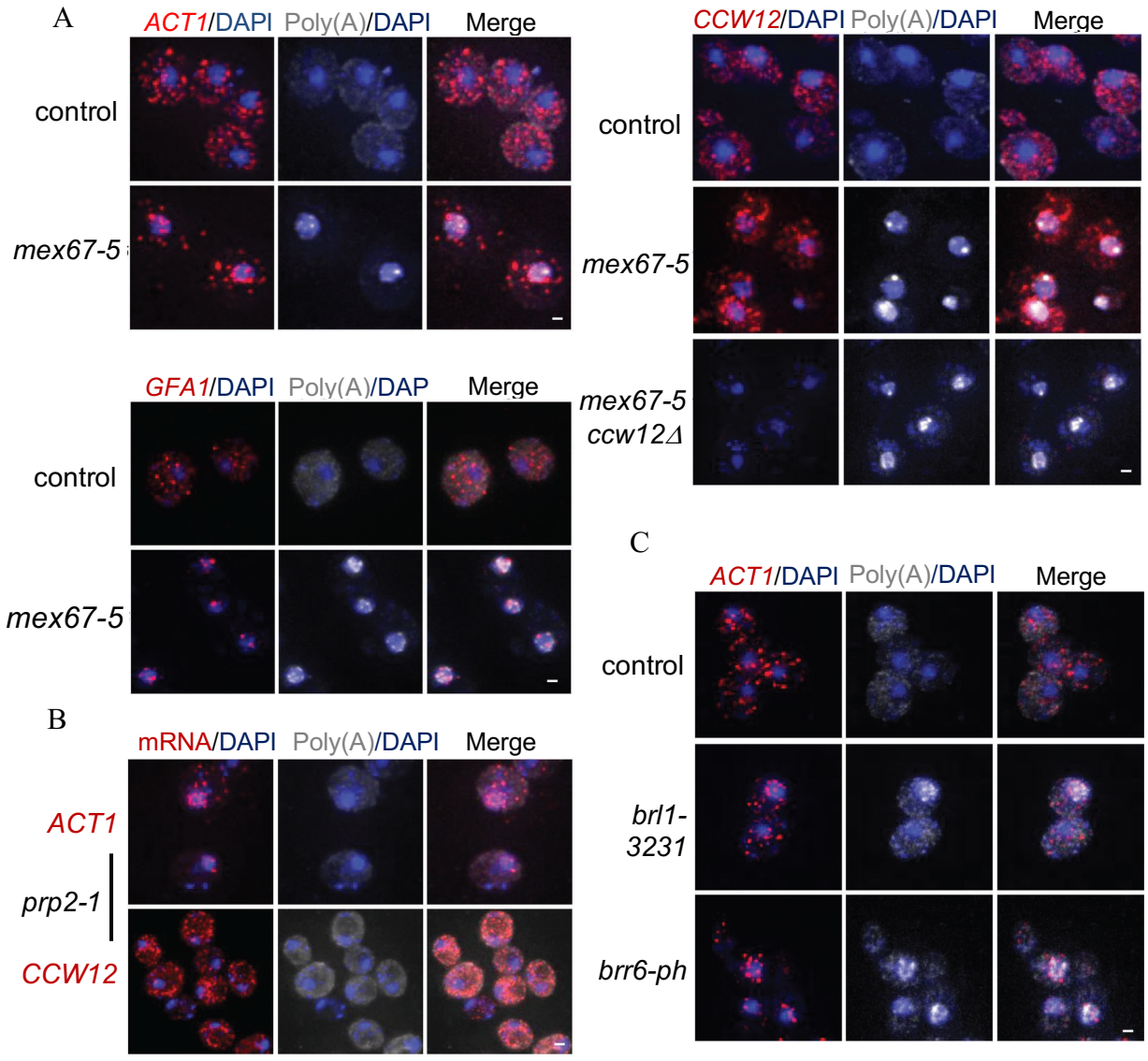


Figure 3-3. Identification of mRNA export mutants using single molecule FISH. (A) Representative images showing *ACT1*, *GFA1*, or *CCW12* mRNA (red) localization in control (*ura10Δ*), *mex67-5,* or *mex67-5 / ccw12Δ* strains compared to poly(A)-RNA (grey) and DAPI (blue) following 3 hours at 37°C. (B) Representative images showing *ACT1* or *CCW12* mRNA (red) localization in the *prp2-1* strain compared to poly(A)-RNA (grey) and DAPI (blue) following 3 hours at 37°C. (C) Representative images showing *ACT1* mRNA (red) localization in control (*ura10Δ*), *brl1-3231* and *brr6-ph* strains compared to poly(A)-RNA (grey) and DAPI (blue) following 6 hours at 37°C. Scale bars = 1 μm.

in the case of *CCW12*, the probes were specific for the transcript being targeted.

Using smFISH data for *GFA1*, we determined the number of transcripts and distribution of these transcripts between the nucleus and cytoplasm. In a haploid control strain (*ura10Δ*), *GFA1* was found to be present at ~5 copies per cell with ~18% of these transcripts being in the nucleus (based on DAPI and *ITS* signals), while the mRNA export mutant *mex67-5* contained ~88% of transcripts in the nucleus (Table 3-1 and Figure 3-3A). Most mutants showed less than a two-fold change in *GFA1* levels with nuclear pools of the mRNA that varied between 16 - 95% (Table 3-1). In the case of *prp2-1*, there was no impact on *GFA1* localization or transcript number, but *ACT1* export was altered (Figure 3-3B), which is consistent with the role of Prp2p in splicing (Lustig *et al.*, 1986). For *brl1-3231* and *brr6-ph,* these mutants showed no mRNA export defect after a 3-hour temperature shift, but given the role of the gene products in nuclear envelope maintenance and NPC biogenesis (de Bruyn Kops and Guthrie, 2001; Saitoh *et al.*, 2005; Hodge *et al.*, 2010; Lone *et al.*, 2015), we expected to observe a defect. As such, we performed a 6-hour temperature shift providing additional time for *Ts* phenotypes to develop, and under these conditions, mRNA export defects were observed for both mutants (Figure 3-3C).

3.2.4 mRNAs localize to distinct subdomains of the nucleus in mutants

smFISH data showed that several mutants had a large increase in the fraction of nuclear mRNAs (e.g. *mex67-5* and *gle1-4*), while others did not (e.g. *rsp5-3* and *sli15-1*). However, in almost all instances the localization of mRNAs appeared distinct in the mutants tested. For example, strains carrying mutations in genes directly linked to the mRNA export process (i.e. *DBP5, GLE1, MEX67*, and *NUP159)* often had mRNAs near the periphery of the DAPI stained DNA mass (see *mex67-5* in Figure 3-3A). This suggests that within these mutants mRNAs accumulated at or near nuclear pore complexes, as previously reported for mutants of *MEX67* (Hurt *et al.*, 2000; Smith *et al.*, 2015) and *DBP5* (Hodge *et al.*, 2011). The *prp2-1* mutant also showed *ACT1* transcripts near the nuclear periphery (Figure 3-3B), which may be related to quality control mechanisms that block export of pre-mRNAs (Galy *et al.*, 2004; Palancade *et al.*, 2005; Iglesias *et al.*, 2010; Hackmann *et al.*, 2014). Using Ndc1p-GFP as a marker of NPCs and the nuclear periphery, we quantified the percentage of transcripts within ~250 nm of the NPC

signal in individual cells (n=50) of select mutants. In control cells, 38 ± 15 % of *GFA1* mRNAs were found within this distance, which increased in mRNA export mutants to 68 ± 35 % (*mex67-5*) and 61 ± 27 % (*dbp5-1*) (Figure 3-4A). The *prp2-1* mutant did not show an increase in peripheral localization of *GFA1* (39 ± 18 %), while *ACT1* increased from 25 ± 12 % in control to 49 ± 17 % in *prp2-1* (Figure 3-4B). These data are consistent with the accumulation or retention of mRNAs near NPCs when splicing or late steps in the mRNA export pathway are disrupted.

A second distinct localization pattern was the accumulation of mRNAs within a nuclear focus. Rarely present in the control strain, 15 mutants showed a >5-fold increase in the frequency of nuclear foci with an intensity ≥10-fold that of single transcripts, including *rsp5-3* (Figure 3-5). In quantifying *GFA1* smFISH data, such foci were counted as single mRNAs, which likely underestimates the number of nuclear transcripts in these mutants and leads to a lower level of nuclear accumulation reported in Table 3-1. Rsp5p is an ubiquitin-ligase that functions in both the cytoplasm and nucleus (Belgareh-Touzé *et al.*, 2008; Kaliszewski and Zoładek, 2008) and has a known role in mRNA biogenesis via modification of the THO/TREX complex (Neumann *et al.*, 2003; Rodriguez *et al.*, 2003; Gwizdek *et al.*, 2005). Given the functions of Rsp5p and other genes with this phenotype, that various mRNA probes show the same defect (i.e. not related to splicing), and the intensity of the smFISH signal, we speculated that these foci represent gene transcription sites. To test this possibility, a LacO array was integrated ~400 bp upstream of the *ACT1* gene, which could be used in combination with LacI-GFP and FISH probes to localize both the *ACT1* gene and mRNA focus. In *rsp5-3*, bright *ACT1* mRNA foci were in close proximity to the *ACT1* gene locus; separated by an average distance of 0.15 ± 0.05 μm, while the *CCW12* mRNA focus was distinct from the *ACT1* gene locus at an average distance of 0.81 ± 0.27 μm (n=50, Figure 3-6A and B). These data are consistent with these bright mRNA foci being at or near transcription sites. We also observed close association between the *ACT1* mRNA and gene loci in mutants linked to chromosome segregation (*spc24-10* and *sli15-3*), as well as *ldb19Δ* (Figure 3-6C). Ldb19p is a regulator of Rsp5p that functions in ubiquitin-dependent receptor endocytosis (Lin *et al*., 2008), but is not known to have a nuclear role. These data match previous reports of mRNAs being retained near transcription sites in THO/TREX complex and mRNA export pathway mutants following heat shock (Hilleren and Parker, 2001; Jensen *et al.*, 2001b; Libri *et al.*, 2002). Moreover, these results demonstrate that the retention of mRNAs at or near transcription sites is a phenotype shared by a set of mutants with diverse cellular functions.

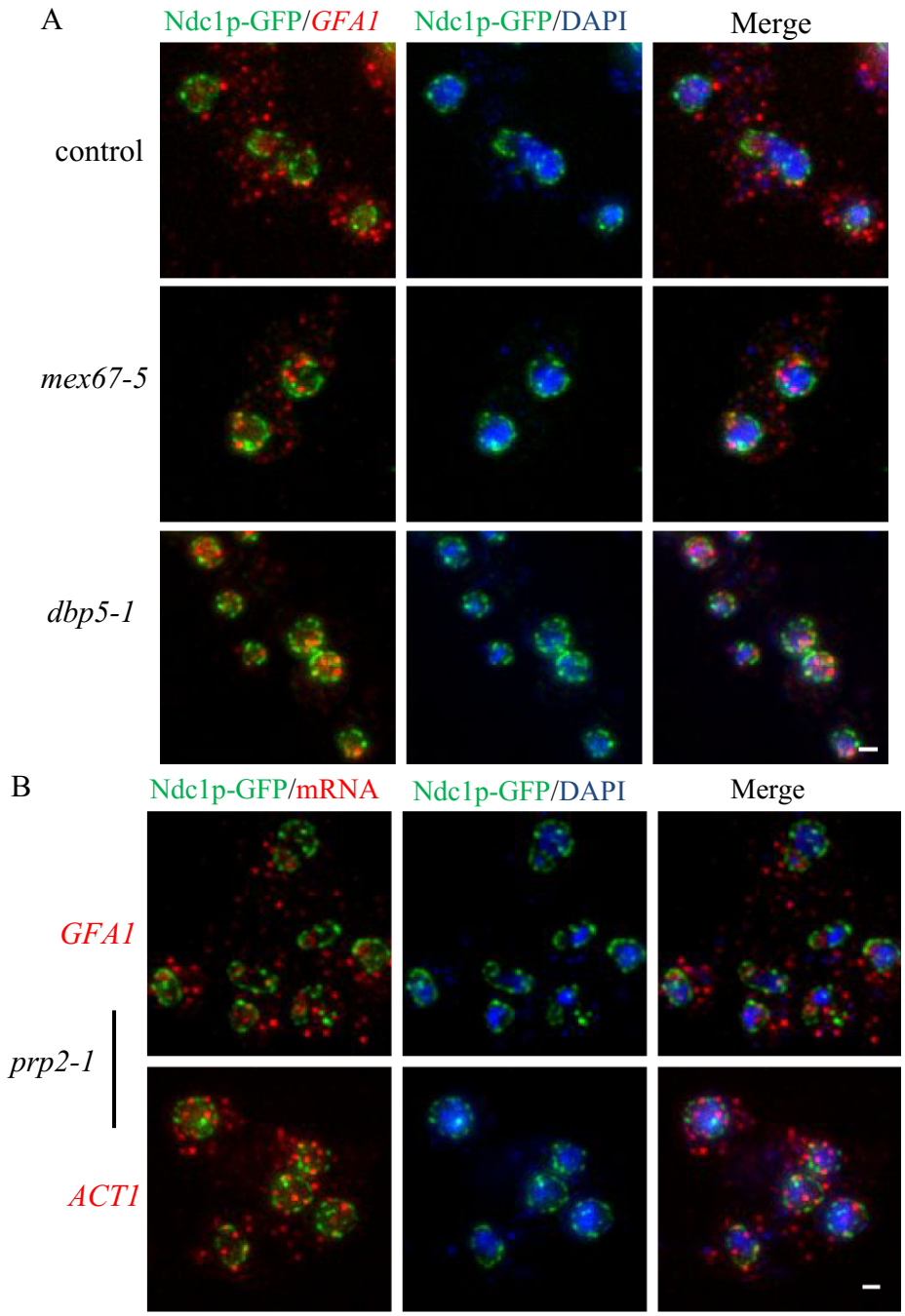


Figure 3-4. Accumulation of mRNAs near the nuclear periphery and NPCs. (A) Representative images showing *GFA1* mRNA (red) localization in control (*ura10Δ*), *dbp5-1, and mex67-5* strains compared to NPCs (green, Ndc1p-GFP) and DAPI (blue) following 3 hours at 37°C**.** (B) Representative images showing *GFA1* or *ACT1* mRNA (red) localization in the *prp2-1* strain compared to NPCs (green, Ndc1p-GFP) and DAPI (blue) following 3 hours at 37°C. Scale bars = 1 μm.

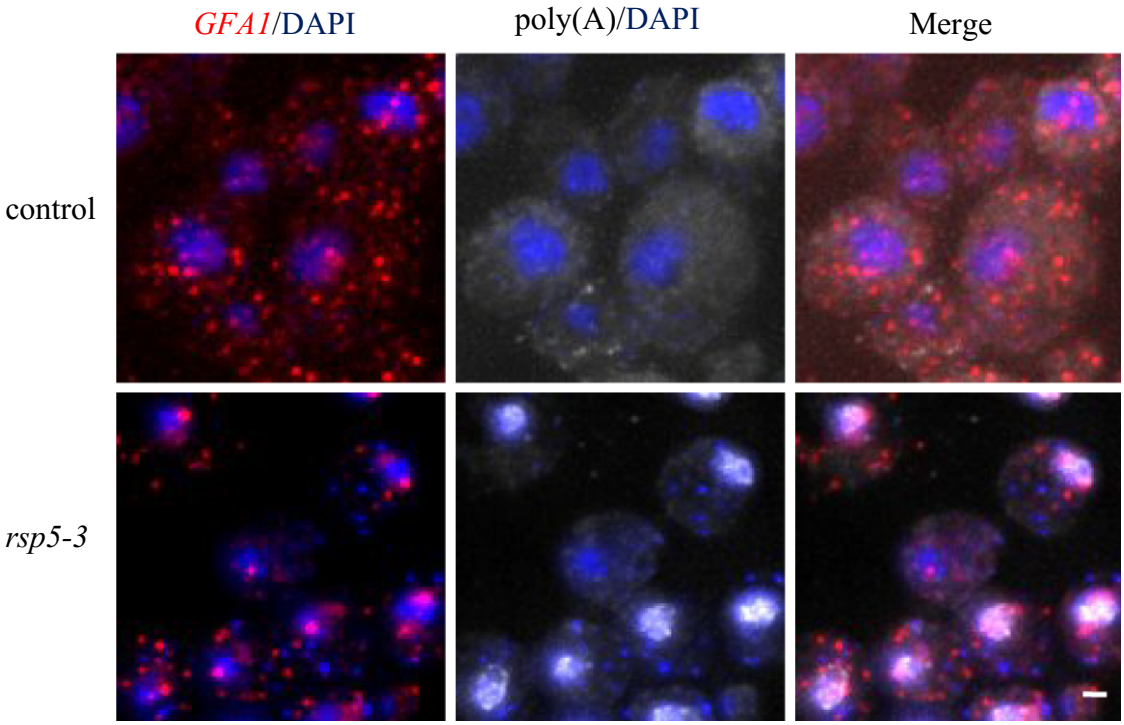


Figure 3-5. Accumulation of *GFA1* mRNA in control and *rsp5-3* strains. Representative images showing *GFA1* mRNA (red) localization in control (*ura10Δ*) and *rsp5-3* strains compared to poly(A)-RNA (grey) and DAPI (blue) following 3 hours at 37°C. Scale bar = 1 μm.

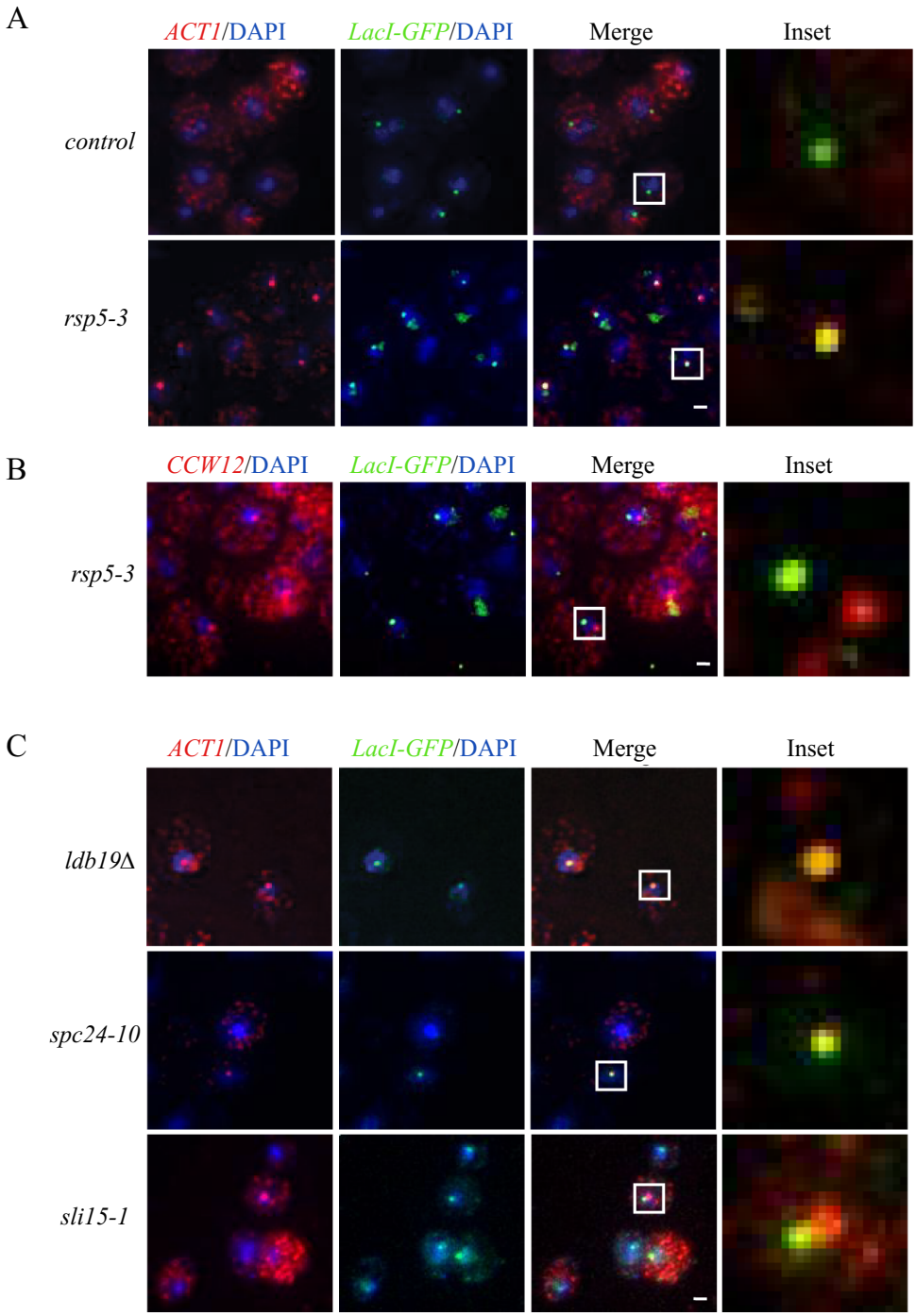


Figure 3-6. Accumulation of mRNAs near transcription sites. (A-B) Representative images showing localization of *ACT1* or *CCW12* mRNA (red) and their respective gene (green) locus in control (*ura10Δ*) and *rsp5-3* strains following 3 hours at 37°C. (C) Representative images showing *ACT1* (red) in *ldb19Δ, spc24-10,* and *sli15-1* strains compared to the *ACT1* gene locus (green, marked by LacO/GFP-LacI) and DAPI (blue) following 3 hours at 37°C. Inset shows a zoomed in view of the boxed region in the merged image. Scale bars = 1 μm.

The third localization pattern was an accumulation of mRNAs next to the DAPI stained DNA mass with or near *ITS* (Table 3-1). This included mutants in components of the exosome (i.e. *DIS3*, *RRP43*, and *CSL4*) and TRAMP complex (*MTR4*), which have previously been reported to accumulate poly(A)-RNA in the nucleolus, as well as the heat shock induced transcript *SSA4* and the localized mRNA *ASH1* (Brodsky and Silver, 2000; Thomsen, 2003; Carneiro *et al.*, 2007; Rougemaille *et al.*, 2007; Du *et al.*, 2008). To quantify nucleolar localization (defined by *ITS* staining), FISH data for *GFA1* was compared between the mRNA export factor (*dbp5-1*) and exosome component (*dis3-1*), which both accumulated *GFA1* transcripts in the nucleus after temperature shift (Figure 3-7A). Using this data, we quantified where nuclear mRNA localized with respect to the *ITS* signal and found 0.4 ± 0.6 and 0.8 ± 0.9 *GFA1* transcripts in the nucleolus of the control (*ura10Δ*) and *dbp5-1* strains. This represents 46% (*ura10Δ*) and 36% (*dbp5-1)* of total nuclear transcripts in these strains (Table 3-1). In contrast, 1.7 ± 1.2 *GFA1* mRNAs were found in the nucleolus of a *dis3-1* strain, which was 80% of the total *GFA1* transcripts in the nucleus of this mutant. This suggests that disruption of exosome function leads to the accumulation of mRNAs within the nucleolus. mRNAs can be classified based on protein binding profiles, which have recently been used to define ten general mRNP classes (Tuck and Tollervey, 2013). These range from mRNAs most likely to be processed and exported to the cytoplasm for translation (Class X) to those that have protein binding patterns similar to CUTs (Class I) that are targeted for nuclear RNA surveillance (Wyers *et al.*, 2005). The three transcripts we observed in the nucleolus (*GFA1*, *ACT1*, and *CCW12*) belong to Class X, so to extend our observations to other mRNA classes, we used gene specific FISH probes to assay the localization of *IMD2* (Class I) and *LEU1* (Class II) mRNAs. In both cases, we observed that these mRNAs were retained in the nucleus of a *dbp5-1* mutant, localized to the nucleolus in *dis3-1*, and appeared as a bright nuclear focus in *rsp5-3* (Figure 3-7B and 3-7C). By using strains carrying a *NAB2-GFP* plasmid and *GFP* FISH probes, we also observed transcripts generated from this plasmid within the nucleolus of a *dis3-1* strain (Figure 3-7D). This implies that mRNAs in the nucleolus are near full length and not short transcripts resulting from early transcription termination since the *GFP* probes are directed against the 3′ end of the transcript. These data demonstrate that various mRNAs localize to distinct subdomains of the nucleus when RNA processing and surveillance pathways are disrupted.

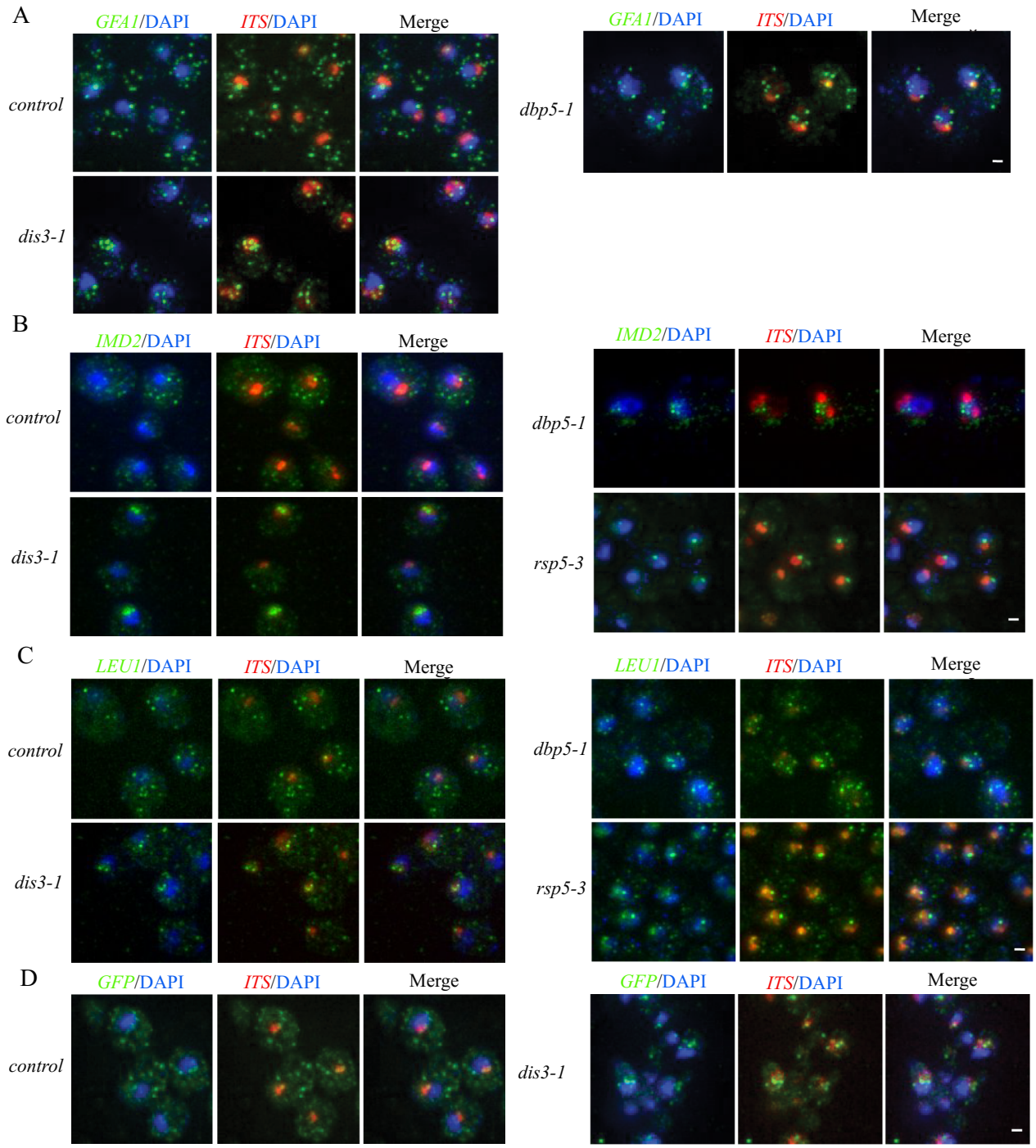


Figure 3-7. Nucleolar localization of mRNA. (A-C) Representative images showing *GFA1*, *IMD2,* or *LEU1* mRNA (green) localization in control (*ura10Δ*), *dbp5-1, dis3-1,* and *rsp5-3* strains compared to *ITS* (red) and DAPI (blue) following 3 hours at 37°C. (D) Representative images showing the localization of *NAB2-GFP* transcripts using *GFP* in situ probes (green) in control (*ura10Δ*) and *dis3-1* strains compared to *ITS* (red) and DAPI (blue) following 3 hours at 37°C. Scale bars = 1 μm.

3.2.5 mRNP-associated factors are sequestered in the nucleolus with mRNA

Following the observation that mRNAs accumulated in the nucleolus, the subcellular localization of three proteins involved in mRNA processing were characterized. Specifically, in control (*ura10Δ*), *dbp5-1*, and *dis3-1* strains we assayed localization of Nab2p-GFP (polyadenosine RNA-binding adaptor protein for Mex67p), Prp19p-GFP (splicing factor), and Hrp1p-GFP (subunit of cleavage factor I; required for the cleavage and polyadenylation of mRNA 3′ ends). In control and *dbp5-1* strains these proteins co-localized with the DAPI stained DNA mass and were largely absent from the nucleolus as marked by *ITS* (Figures 3-8, 3-9, 3-10). In contrast, within the *dis3-1* strain these factors were found throughout the nuclear volume within both the DAPI and *ITS* stained regions, suggesting that factors involved in mRNA biogenesis redistribute to the nucleolus in this mutant, similar to mRNAs. We further tested if loss of Rrp6p activity would lead to the re-localization of these same factors. Rrp6p is a non-essential catalytic subunit of the exosome that when mutated results in poly(A)-RNA accumulation in a discrete domain within the nucleolus (Hieronymus *et al.*, 2004; Carneiro *et al.*, 2007; Rougemaille *et al.*, 2007). In a *rrp6Δ* strain we observed ~5 *GFA1* transcripts per cell with 18% being nuclear, which is comparable to the control strain (Table 3-1), but 71% of these nuclear transcripts co-localized with *ITS* as compared to 46% in control. Nab2p-GFP, Prp19p-GFP, and Hrp1p-GFP were also enriched within the nucleolus (Figures 3-8, 3-9, 3-10), demonstrating that both mRNA and mRNP-associated factors are redistributed within the nucleus of a *rrp6Δ* strain.

In addition to mutations in exosome or TRAMP complex components, mutations in *ENP1* and *SRM1* caused accumulation of poly(A)-RNA next to the DAPI stained DNA mass (Figure 3-11A and Table 3-1). Enp1p functions in pre-rRNA processing and 40S subunit synthesis and Srm1p facilitates nucleocytoplasmic trafficking (Tachibana *et al.*, 1994; Koepp *et al.*, 1996; Chen *et al.*, 2003). *GFA1* mRNAs could be readily observed with the poly(A)-RNA signal that is adjacent to DAPI in both *enp1-1* and *srm1-ts* strains (Figure 3-10A), but *ITS* expression was severely reduced in both mutants preventing us from quantifying *GFA1* localization. Nab2p-GFP, Prp19p-GFP, and Hrp1p-GFP were also localized to the nucleolus of the *enp1-1* strain (Figure 3-11B), but *srm1-ts* could not be characterized due to defects in protein import. These findings demonstrate that localization of mRNAs and associated proteins to the nucleolus does not only

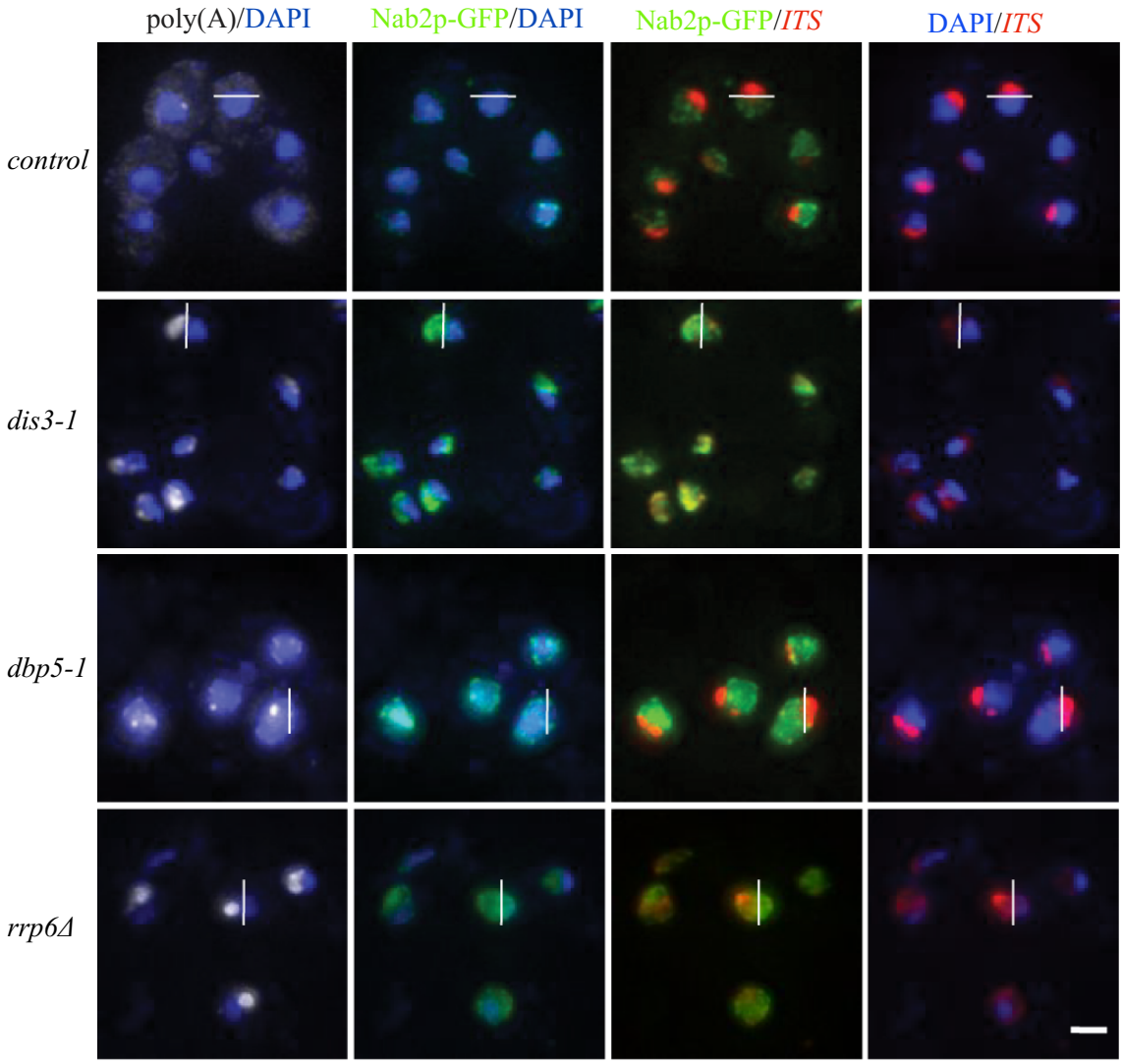


Figure 3-8. Nucleolar enrichment of Nab2p-GFP in exosome mutants. Representative images showing localization of Nab2p-GFP (green) in control (*ura10Δ*), *dbp5-1, dis3-1*, or *rrp6Δ* strains compared with poly(A)-RNA (gray), *ITS* (red), and DAPI (blue) after 3 h at 37°C. A white line has been added to one cell in each image to denote the border between DAPI and *ITS* signals and aid in comparisons. Scale bars = 2 μm.

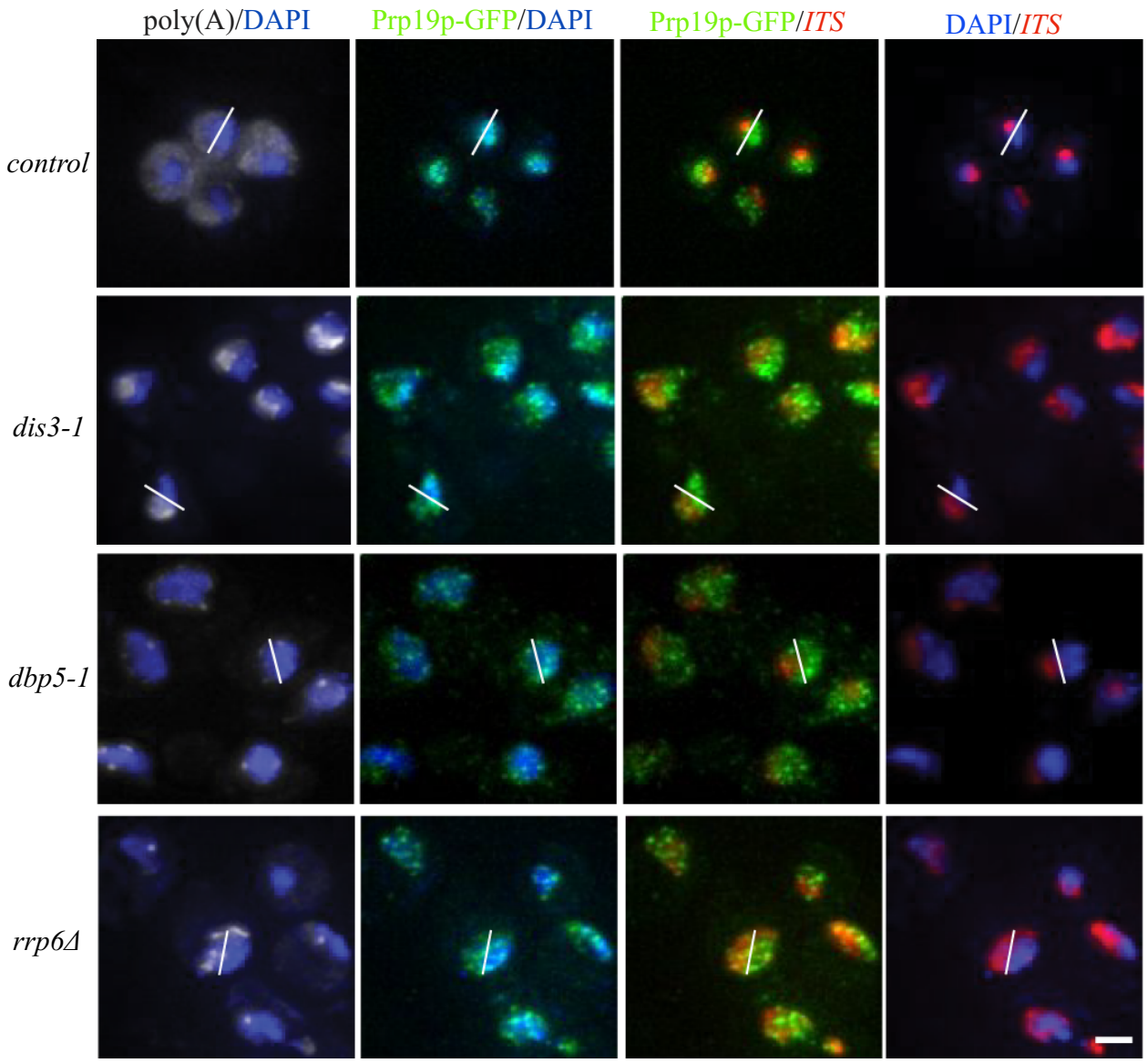


Figure 3-9. Nucleolar enrichment of Prp19-GFP in exosome mutants. Representative images showing localization of Prp19p-GFP (green) in control (*ura10Δ*), *dbp5-1, dis3-1*, or *rrp6Δ* strains compared with poly(A)-RNA (gray), *ITS* (red), and DAPI (blue) after 3 h at 37°C. A white line has been added to one cell in each image to denote the border between DAPI and *ITS* signals and aid in comparisons. Scale bars, 2 μm.

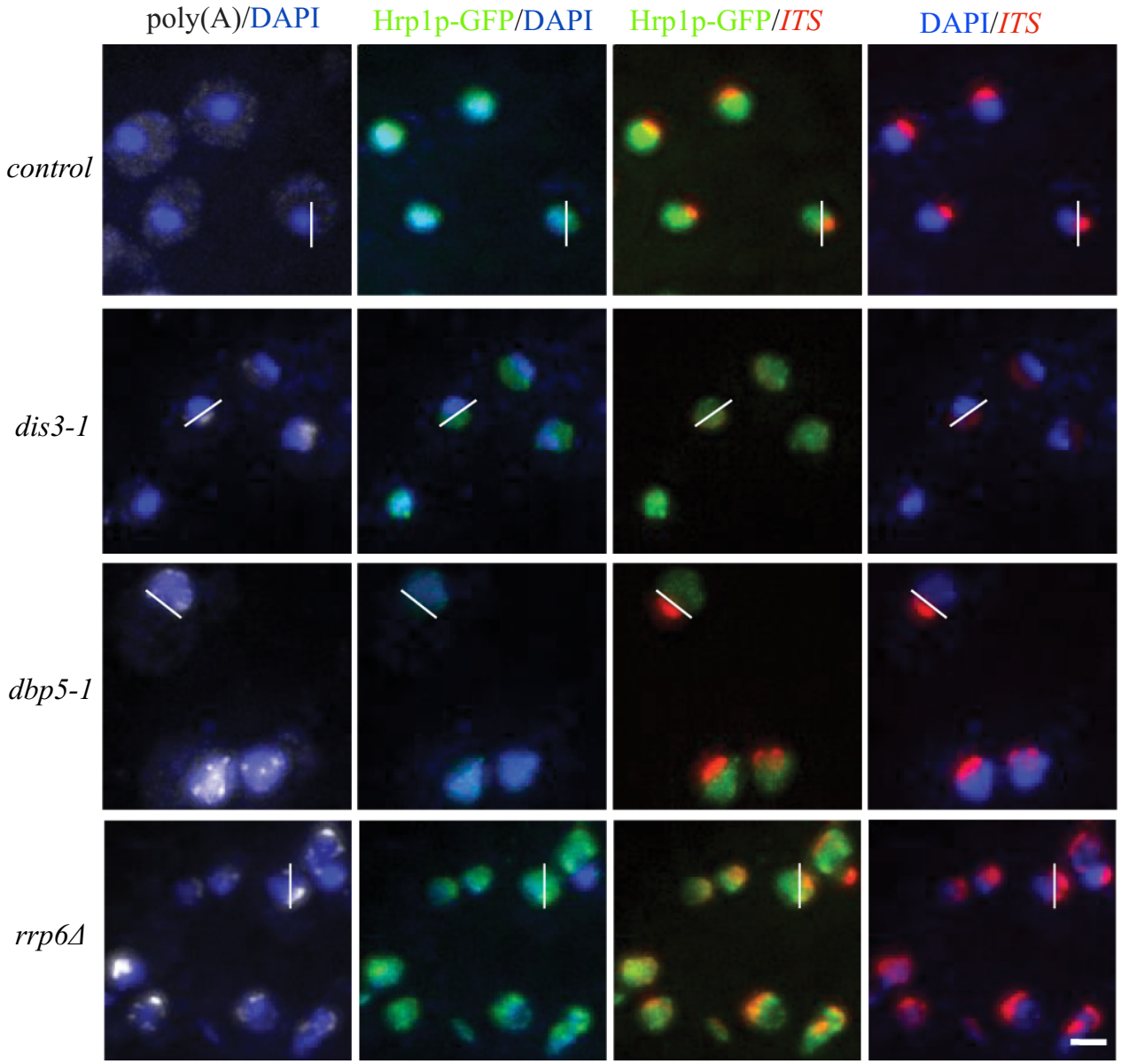


Figure 3-10. Nucleolar enrichment of Hrp1p-GFP in exosome mutants. Representative images showing localization of Hrp1p-GFP (green) in control (*ura10Δ*), *dbp5-1, dis3-1*, or *rrp6Δ* strains compared with poly(A)-RNA (gray), *ITS* (red), and DAPI (blue) after 3 h at 37°C. A white line has been added to one cell in each image to denote the border between DAPI and *ITS* signals and aid in comparisons. Scale bars = 2 μm.

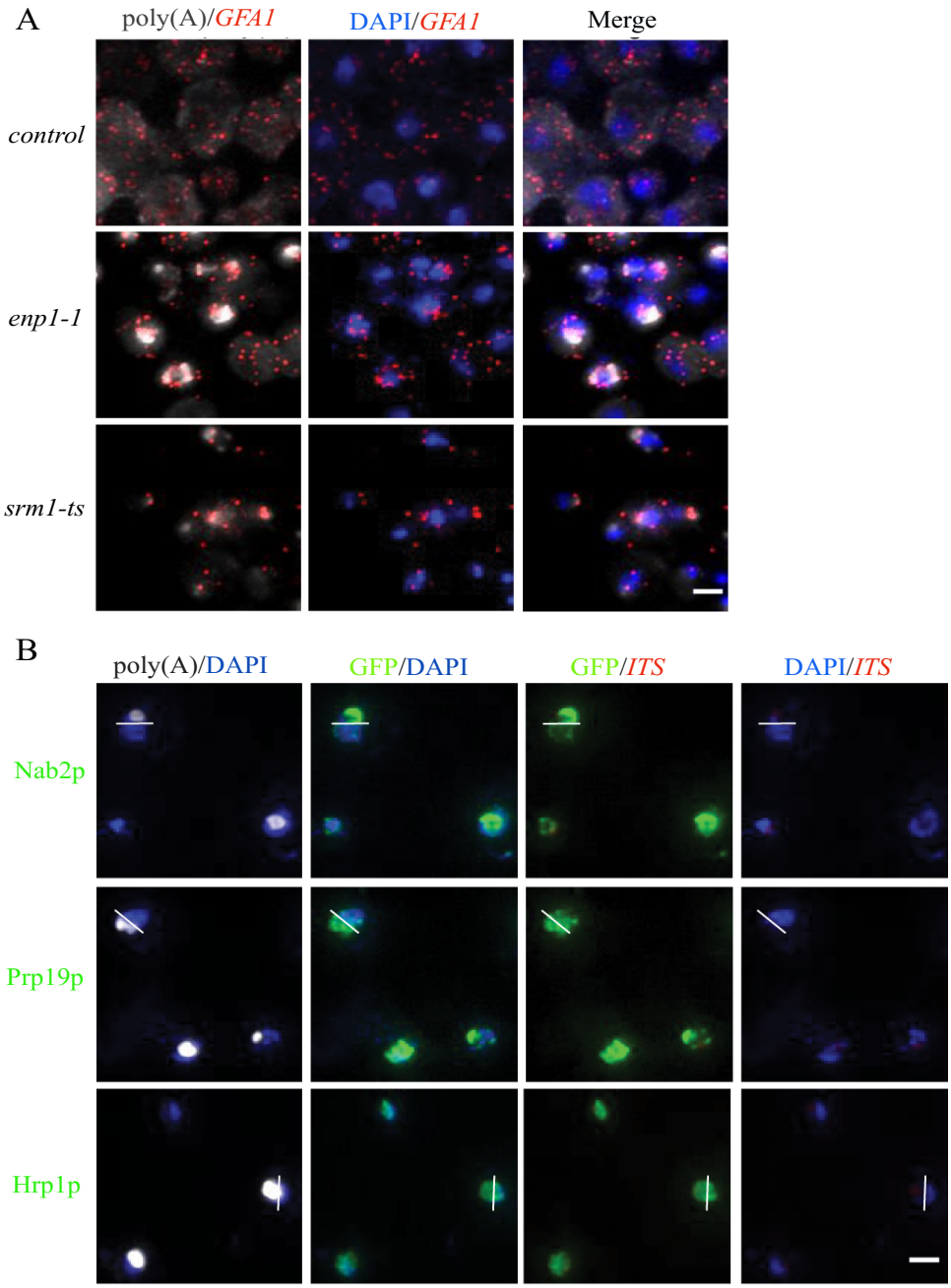


Figure 3-11. Nucleolar enrichment of *GFA1* mRNA and mRNP-associated proteins in *enp1-1*. (A) Representative images showing *GFA1* mRNA (red) localization in control (*ura10Δ*), *enp1-1*, and *srm1-ts* strains compared with poly(A)-RNA (gray) and DAPI (blue) after 3 h at 37°C. (B) Representative images showing localization of Nab2p-GFP, Prp19p-GFP, or Hrp1p-GFP (green) in the *enp1-1* strain compared with poly(A)-RNA (gray), *ITS* (red), and DAPI (blue) after 3 h at 37°C. A white line has been added to one cell in each image to denote the border between DAPI and *ITS* signals and aid in comparisons. Scale bars = 1 μm.

occur when RNA surveillance machinery is mutated and can occur as a result of disruption to other nuclear processes.

3.3 Discussion

mRNA biogenesis involves the interaction of numerous protein factors with each mRNA in a spatially and temporally regulated manner within the nucleus to link the processes of transcription, capping, splicing, 3′ end formation, RNA surveillance, and NPC-mediated export (Oeffinger and Zenklusen, 2012; Niño *et al.*, 2013). Nuclear events also act to define the fate of each mRNP, which can include marking mRNPs in the nucleus for translation, storage, transport, or decay within the cytoplasm (Müller-McNicoll and Neugebauer, 2013; Singh *et al.*, 2015). Consequently, the proper and efficient coordination of these nuclear events is central to the fidelity of the gene expression program with previous works having identified factors involved in these processes through analysis of individual mutants, genetic screens, and comprehensive screening of nonessential genes (Shiokawa and Pogo, 1974; Amberg *et al.*, 1992; Kadowaki *et al.*, 1992, 1994b; Doye *et al.*, 1994; Fabre *et al.*, 1994; Gorsch *et al.*, 1995; Hieronymus *et al.*, 2004). However, the critical nature of these events dictates that essential genes would likely be involved in this process, which prompted us to perform a screen of essential genes for function(s) related to mRNA biogenesis in *S. cerevisiae*. This led to the identification of 29 genes with a nuclear poly(A)-RNA accumulation phenotype (Table 3-1); of which approximately half (15/29) had not previously been reported to have such a defect. Of these, 26 genes were subsequently found to alter mRNA biogenesis and export when mutated based on the observed accumulation of mRNA in the nucleus and/or altered localization of nuclear mRNAs.

3.3.1 Classes of mRNA biogenesis and export mutants

Gene ontology places the majority of the 29 genes into three biological processes: RNA export from nucleus, nuclear RNA catabolic processes, and chromosome segregation. The identification of mRNA export factors is expected and validates our screen. Links between

mRNA export and RNA decay are also well established (Porrua and Libri, 2013; Eberle and Visa, 2014), with various components of the decay machinery having originally been identified in screens for mRNA transport (MTR) mutants (Kadowaki *et al.*, 1992, 1994b). Beyond the related biological functions of the gene products, 26 of the 29 mutants could be further classified into three distinct groups (referred to as class A, B or C) based on shared phenotypes with respect to the localization of poly(A)-RNA and mRNA within the nucleus.

Class A mutants are represented by genes directly involved in mRNA export and NPC function (e.g. *DBP5* and *NUP159*) and were characterized by a diffuse nuclear poly(A)-RNA signal (+/- discrete foci) overlapping the DAPI stain often accompanied by disrupted nucleoli (Figure 3-2 and Table 3-1). These mutants further displayed an increased proportion of mRNAs near the nuclear periphery and NPCs (Figure 3-4). This suggests that within class A mutants mRNPs accumulated at or near nuclear pore complexes due to failures in NPC-mediated export, which is consistent with the known function of these factors. In support of the smFISH observations and data interpretation, delays in export at the nuclear periphery have been observed using live cell imaging techniques for mutants of *DBP5*, *MEX67*, NAB2, and the Nups *MLP1/2* (Hodge *et al.*, 2011; Saroufim *et al.*, 2015; Smith *et al.*, 2015). Of the class A mutants, *prp2-1* showed a transcript specific block of mRNA export with the *ACT1* mRNA localized near the nuclear periphery, but not another mRNA that lacked an intron (Figures 3-3B and 3-4B). This observation is consistent with the description of NPC-dependent quality control mechanisms that prevent the export of immature mRNAs that include pre-mRNAs (Galy *et al.*, 2004; Palancade *et al.*, 2005; Iglesias *et al.*, 2010; Hackmann *et al.*, 2014).

Class B mutants, similar to class A, showed a diffuse poly(A)-RNA distribution (+/- discrete foci) overlapping the DAPI stain (see *rsp5-3*, Figure 3-2), but differed in that mRNAs were retained at or near gene transcription sites (Figure 3-5). Class B mutants included *RSP5, PTA1,* and *CLP1* that are each known to be involved in 3′ end processing (Minvielle-Sebastia *et al.*, 1997; Preker *et al.*, 1997; Zhao *et al.*, 1999; Gwizdek *et al.*, 2005) and could be expected to retain mRNAs at transcription sites due to defects in mRNA biogenesis similar to other mutants affecting co-transcriptional processing (Hilleren *et al.*, 2001; Jensen *et al.*, 2001a; Libri *et al.*, 2002; Zenklusen *et al.*, 2002; Thomsen *et al.*, 2008). Ldb19p is a known regulator of Rsp5p in endocytosis (Lin *et al.*, 2008) that was also found to accumulate mRNAs near transcription sites when mutated, like *rsp5-3* (Figure 3-5). The functional relationship between Ldb19p and Rsp5p

with the shared mRNA processing defects suggests that Ldb19p also regulates the nuclear function of Rsp5p. In support of a nuclear role for Ldb19p, we were able to detect GFP-Ldb19p in the nucleus when overexpressed from a GAL promoter in a *xpo1-1* mutant that disrupts nucleocytoplasmic transport (Figure 3-12).

Unexpectedly, the majority of class B mutants isolated in our screen (9 of 15) were linked to chromosome segregation and cell division. A previous report had noted that mutants affecting chromosome segregation in *S. cerevisiae* displayed poly(A)-RNA accumulation only when the cell undergoes continuous cell division. (Cole *et al.*, 2002). Given the requirement for ongoing cell division, the large number of chromosome segregation mutants with a class B phenotype, and that these mutants show defects in <25% of cells after 3 hours at non-permissive temperature (Table 3-1), it is likely that these defects are related to improper chromosome segregation. This could occur through the random loss of genetic material that encodes any one of the many factors involved in mRNA biogenesis, which would be expected to give rise to variable phenotypes depending on the processing event affected. However, when poly(A)-RNA accumulated in these mutants it is often associated with mRNAs appearing at transcription sites and disruption of the nucleolus (Figures 3-2). A possible explanation for this more constant phenotype is that specific chromosomes or regions of DNA are repeatedly lost in these mutants and this gives rise to the class B defect due to the specific gene(s) being affected. Alternatively, it has been reported that aneuploid states involving different chromosome imbalances in *S. cerevisiae* result in a set of common cellular characteristics (Torres *et al.*, 2007; Sheltzer *et al.*, 2012), which class B mRNA export phenotypes could exemplify. While the molecular details remain unknown, a recent screen for mutants that alter tRNA processing (Wu *et al.*, 2015) did not report a defect for any chromosome segregation mutants and only had two genes (*NUP133* and *RSP5*) in common with the class B mutants identified here. This lack of overlap hints at a more direct relationship between mRNA export and chromosome segregation, as it is expected that both mRNA and tRNA processing would be equally susceptible to disruption by random chromosome segregation errors.

Finally, class C mutants displayed poly(A)-RNA accumulation next to the DAPI stain with or adjacent to the nucleolus (based on *ITS*) and an increased frequency of mRNAs within this compartment (Figures 3-2, 3-5, and 3-6). Many of these mutants are involved in RNA processing and surveillance as components of the exosome and TRAMP complexes, which raises

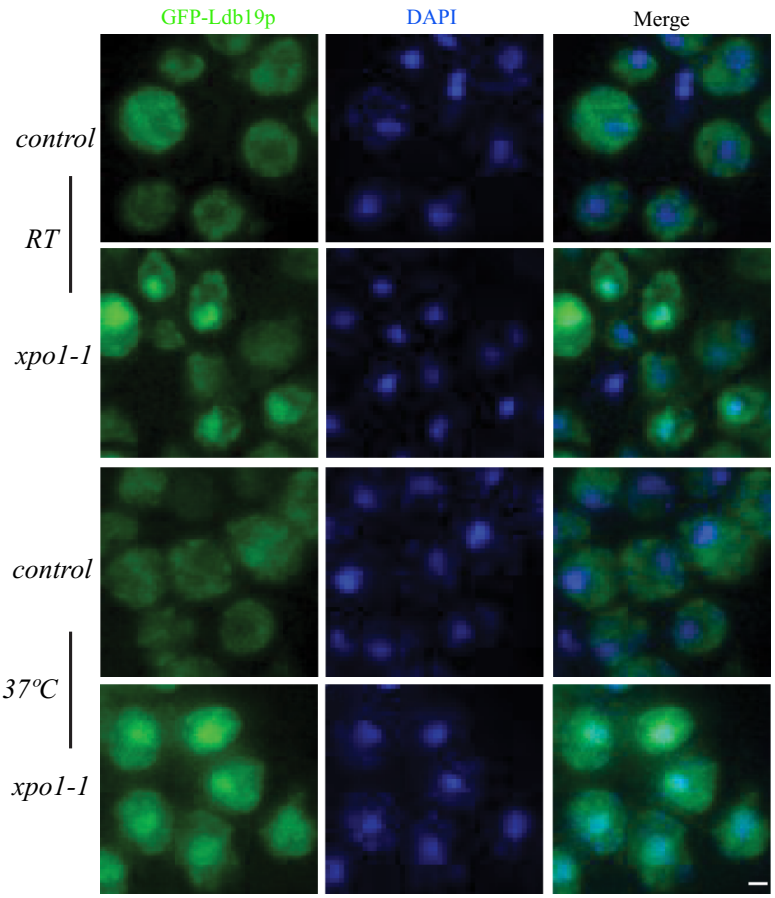


Figure 3-12. Nuclear localization of Ldb19p-GFP. Representative images showing GFP-Ldb19p (green) and DAPI staining (blue) in control (*ura10Δ*) and *xpo1-1* strains at room temperature and after shift to 37°C for 1 hour. Cells were grown in galactose media overnight to induce GFP-Ldbp19p expression. Scale bar = 1 μm.

the possibility that the detected mRNAs are aberrant and only become detectable due to a loss of surveillance and decay activities. Yet, we did not detect an increase in the number of *GFA1* transcripts in these mutants (Table 3-1), which may be expected if these were common products of transcription that are normally degraded. By using probes against the *GFP* coding sequence at the 3′ end of an mRNA (Figures 3-6), we also observed nucleolar mRNA accumulation suggesting that these transcripts are near full length and not the product of early transcript termination events. Other mutants (i.e. *enp1-1* and *srm1-ts*) that are not part of the decay machinery caused similar mRNA localization phenotypes suggesting that nucleolar mRNAs could be regulated by mechanism other than those attributable to RNA surveillance machinery.

In addition to mRNAs, we also observed mRNP-associated factors enriched within the nucleolus in class C mutants (Figures 3-8, 9, 10). The redistribution of mRNP-associated factors may occur as constituents of mRNPs present within the nucleolus, but it is also possible that poly(A)-RNA binding proteins, such as Nab2p (Anderson *et al.*, 1993; Green *et al.*, 2002), localize to the nucleolus due to promiscuous binding of the protein to accumulated poly(A)-RNA. Thus, the inappropriate localization of poly(A)-RNA binding proteins, which include mRNA export adaptors, could explain why class C mutants with various cellular functions cause the same terminal phenotype. The heat shock induced transcript *SSA4*, the localized mRNA *ASH1*, and *PHO84* antisense RNA have also been reported to enter the nucleolus (Brodsky and Silver, 2000; Thomsen, 2003; Carneiro *et al.*, 2007; Du *et al.*, 2008; Castelnuovo *et al.*, 2013) and mRNAs have recently been observed to transit through the nucleolus for export in live cells (Saroufim *et al.*, 2015). As such, further study is needed to address how often mRNAs enter the nucleolus during normal mRNA biogenesis, the functional significance of this event, and the molecular mechanism(s) facilitating nucleolar localization.

3.3.2 Nuclear homeostasis and mRNA biogenesis

Disruption of different nuclear processes resulted in three classes (referred to as Class A, B, and C, see discussion above) of mRNA biogenesis defects. We envision this occurring by multiple mechanisms: (1) directly as a result of mutation in a factor involved in mRNA biogenesis and export, (2) by activating quality control mechanisms that retain immature/aberrant

mRNPs, and (3) due to changes in nuclear structure/function that result in inefficiencies and failures in mRNA biogenesis and export. The first two mechanisms are exemplified by the *mex67-5* mutation, as it is well established that Mex67p functions directly in mRNA export at NPCs (Segref *et al.*, 1997; Santos-Rosa *et al.*, 1998; Hurt *et al.*, 2000; Sträßer *et al.*, 2000; Lund and Guthrie, 2005; Smith *et al.*, 2015) and a quality control mechanism was recently described that protects cells during loss of Mex67p function by retaining mRNAs at transcription sites (Kallehauge *et al.*, 2012). Moreover, both mechanisms are highlighted by our smFISH data where we observed mRNAs near the nuclear periphery (i.e. export failure) and ~8 fold increase in transcription site foci (i.e. retention) in the *mex67-5* strain. Similar quality control mechanisms have been described in other mutants to cause mRNAs to be retained near transcription sites and at NPCs to prevent export of pre-mRNAs and aberrant mRNPs (Hilleren *et al.*, 2001; Libri *et al.*, 2002; Galy *et al.*, 2004; Palancade *et al.*, 2005; Rougemaille *et al.*, 2008; Saguez *et al.*, 2008; Iglesias *et al.*, 2010; Hackmann *et al.*, 2014), which together may explain many of the observed localization patterns in class A and B mutants.

The third mechanism leading to mRNA biogenesis and export defects likely operates at the systems level and is represented by *srm1-ts*, *enp1-1*, and the many mutants that affect chromosome segregation and exosome function. In these cases, altering nucleocytoplasmic transport, various aspects of RNA biogenesis, and inducing aneuploidy would alter nuclear homeostasis, and in turn, disturb mRNA biogenesis and export. For example, the *enp1-1* mutation induced strong poly(A)-RNA accumulation in the nucleolus and redistribution of mRNA processing factors to this compartment (Figures 3-7, 8, 9, and 10), which likely impacts mRNA processing and export by making these essential factors limiting. This relationship has also been observed to work in reverse with mRNA export defects causing nucleolar disruption (Kadowaki *et al.*, 1994c; Dockendorff *et al.*, 1997; Segref *et al.*, 1997; Thomsen *et al.*, 2008). In this way, *enp1-1* (and other such mutants) function by perturbing overall nuclear homeostasis and would not be considered an mRNA processing or export factor per se.

Within this paradigm of system level perturbations, how exactly chromosome segregation mutants impact the cell and cause specific mRNA processing and export defects remain to be determined. Still, we find it noteworthy that aneuploidy is known to alter tumorigenesis through mechanisms that include changes in gene expression (Gordon *et al.*, 2012; Holland and Cleveland, 2012; Santaguida and Amon, 2015; Dürrbaum and Storchová, 2016) and our work

now highlights the fact that many chromosome segregation mutants cause the accumulation of mRNAs at transcription sites, which provides one mechanism by which this may occur. While these types of relationships could be casually termed as indirect, we would argue that these system level perturbations are important to understand since it is likely that disease states arise from mutation(s) due to both the specific process impacted by the mutation and the system level changes that are induced.

Chapter IV: *Stabilization of polyadenylated noncoding RNA species leads to a generalized disruption in nuclear RNA homeostasis*

---

* A version of this chapter is being prepared for submission as *Paul, B., Aguilar, LC., Gendron, L., Rajan, AAR., Montpetit, R., Trahan, C., Pechmann, S., Oeffinger, M. and Montpetit. B. Stabilization of polyadenylated noncoding RNA species leads to an overall disruption in nuclear RNA homeostasis.*

4.1 Introduction

In *S. cerevisiae* and other eukaryotes, both ncRNAs (e.g. snoRNAs and rRNAs) and mRNAs are synthesized and processed in the nucleus as a part of the gene expression program. Each transcript dynamically associates with a set of RBPs as part of an RNP, with the associated RBPs guiding processing events (Matera *et al.*, 2007; Singh *et al.*, 2015; Kressler *et al.*, 2017). During biogenesis, RNPs are also actively surveyed by quality control mechanisms to recognize and decay aberrantly processed mRNAs and ncRNAs, such that these transcripts do not continue along the gene expression pathway and synthesize aberrant proteins or RNA (Fasken and Corbett, 2009; Ghosh and Jacobson, 2010; Hopper and Huang, 2015; Kressler *et al.*, 2017). Consequently, competition between biogenesis and surveillance factors for the maturing transcript within the nucleus is likely critical for determining transcript fate (i.e. decay vs. maturation and export) and shaping the overall cellular transcriptome (Porrua and Libri, 2013; Bresson and Tollervey, 2018).

The exosome constitutes a central part of the gene expression system that functions to maintain cellular RNA homeostasis (Chlebowski *et al.*, 2013; Porrua and Libri, 2013; Schneider and Tollervey, 2013). At its core, the exosome is a 10 subunit protein complex that harbours RNA binding and nuclease activities that are used to carry out RNA biogenesis and quality control, including Dis3p/Rrp44p, which possesses both exo- and endonuclease activities (Mitchell *et al.*, 1997; Allmang *et al.*, 1999b). In addition, the nuclear exosome contains an eleventh subunit, Rrp6p, which functions as a 3′-5′ exonuclease (Briggs *et al.*, 1998; Burkard and Butler, 2000; van Hoof *et al.*, 2000a). The core subunits of the exosome form a barrel (6 subunits) and cap-like structure (3 subunits) that together promote unwinding of the RNA and subsequent cleavage by the two catalytic subunits, Dis3p/Rrp44p and Rrp6p (Liu *et al.*, 2006; Wang *et al.*, 2007; Bonneau *et al.*, 2009; Makino *et al.*, 2013). Csl4p, a component of the cap structure, facilitates substrate recognition (van Hoof *et al.*, 2000b; Lorentzen *et al.*, 2007) and recruitment of Rrp6p to the exosome (Kowalinski *et al.*, 2016; Zinder *et al.*, 2016). In addition, the Trf4p/Trf5p-Air1p/Air2p-Mtr4p polyadenylation (TRAMP) and Nrd1p-Nab3p-Sen1p (NSS) complexes support exosome-mediated RNA processing and decay activities by directing RNA substrates to the exosome (de la Cruz *et al.*, 1998; Steinmetz *et al.*, 2001; LaCava *et al.*, 2005; Vanácová *et al.*, 2005; Wyers *et al.*, 2005; Vasiljeva and Buratowski, 2006).

Through its nuclease-associated activities, and in conjunction with the TRAMP and NNS complexes, the exosome participates in both RNA biogenesis and surveillance of different types of stable RNAs, such as rRNAs, snoRNAs, and mRNAs (Chlebowski *et al.*, 2013; Schmidt and Butler, 2013; Schneider and Tollervey, 2013; Porrua and Libri, 2015). Consequently, when exosome activity is lost, or TRAMP or NNS functions are abolished, the biogenesis and surveillance of these RNA species are severely perturbed resulting in the stabilization of various RNA processing intermediates (Steinmetz *et al.*, 2001; LaCava *et al.*, 2005; Wyers *et al.*, 2005; Vasiljeva and Buratowski, 2006; Gudipati *et al.*, 2012; Schneider *et al.*, 2012; Tudek *et al.*, 2014; Choque *et al.*, 2018). For example, in exosome, TRAMP, and NNS mutants polyadenylated rRNA species and 3′-extended polyadenylated snoRNAs have been shown to accumulate due to stalled or defective processing events, with this material in some instances localizing to discrete subdomains of the nucleus (van Hoof *et al.*, 2000a; LaCava *et al.*, 2005; Dez *et al.*, 2006; Houseley and Tollervey, 2006; Kadaba *et al.*, 2006; Carneiro *et al.*, 2007; Grzechnik and Kufel, 2008; Wery *et al.*, 2009; Leporé and Lafontaine, 2011).

Loss of nuclear RNA surveillance activity further leads to unstable products of transcription to be stabilized that are often at low to non-detectable levels in wild-type cells. Collectively, these are referred to as pervasive transcripts, but based on the mutant background in which they were originally detected, they have been classified into various groups: cryptic unstable transcripts (CUTs) and stable unannotated transcripts (SUTs) in a *rrp6Δ* mutant (Xu *et al.*, 2009; Vera and Dowell, 2016), and Nrd1 unterminated transcripts (NUTs) in a *nrd1Δ* mutant (Schulz *et al.*, 2013). In addition, loss of another exonuclease involved in RNA processing and decay, Xrn1p, results in the stabilization of a related class of transcripts termed XUTs (Van Dijk *et al.*, 2011). Both CUTs and NUTs undergo transcriptional termination via the NNS complex in a manner similar to that by which the NNS complex functions in snRNA and snoRNA processing (Steinmetz *et al.*, 2001; Arigo *et al.*, 2006b; Thiebaut *et al.*, 2006; Schulz *et al.*, 2013; Tudek *et al.*, 2014). However, most SUTs and XUTs are terminated by the cleavage and polyadenylation factor (CPF) pathway, with SUTs being degraded through activities of the exosome, NMD pathway, and Xrn1p, while XUTs are generally exported and degraded in the cytoplasm by Xrn1p (Marquardt *et al.*, 2011; Van Dijk *et al.*, 2011).

This highlights the fact that in many instances the biogenesis and decay of different classes of RNAs are facilitated by an overlapping set of machineries (e.g. exosome, NNS, and

TRAMP complex), including a number of shared RBPs. Recent studies described RBP binding patterns across the yeast and human cellular transcriptome and highlighted similarities and differences in RBP binding to different RNA classes (Tuck and Tollervey, 2013; Sohrabi-Jahromi *et al.*, 2019). A key finding is that particular RBPs, including Nab2p and Hrp1p, function in both the biogenesis of mRNAs and surveillance of pervasive transcripts (Tuck and Tollervey, 2013). These discrete functions, in the context of different transcript classes, would suggest that the amount of Nab2p and Hrp1p available to function in one pathway (i.e. mRNA biogenesis) would be impacted by the amount of that RBP engaged in the other process (i.e. surveillance and decay). As such, imbalances in processing and surveillance that lead to the sequestration of an RBP within one of these pathways could potentially result in a wider dysregulation of gene expression by limiting the activity of the RBP in other pathways. Examples of this may be observed in exosome and TRAMP mutants, which cause poly(A)-RNA accumulation as well as sequestration of Nab2p and Hrp1p within the nucleolus (Carneiro *et al.*, 2007; Paul and Montpetit, 2016).

The accumulation of the poly(A)-RNA binding protein (PABPs) Nab2p and Hrp1p in the nucleolus of select RNA processing mutants (Carneiro *et al.*, 2007; Paul and Montpetit, 2016) is of particular interest given the emerging literature linking PABPs, mRNA export, and nuclear decay. For example, upon loss of Nab2p activity there is an observed increase in nuclear decay of newly transcribed transcripts (Tudek *et al.*, 2018b). In contrast, Nab2p overexpression *in vivo* (Tudek *et al.*, 2018b), or presence *in vitro* (Schmid *et al.*, 2015), suppresses mRNA decay. These findings suggest a competition model for the poly(A) tail, in which PABPs bind and protect the transcript during nuclear maturation to promote export, while decay factors compete with PABPs for the poly(A) tail for the purpose of promoting decay (Porrua and Libri, 2013; Tudek *et al.*, 2018b, 2019). Hence, alterations in RBP availability would be expected to feedback on nuclear RNA homeostasis due to this competitive relationship, with variations in the level of nuclear poly(A)-RNA acting as a mechanism by which RBP availability could be modulated with a direct impact on the cellular transcriptome.

Disruption of either *ENP1* or *SRM1*, two genes that are not known to be involved in RNA surveillance and decay, caused phenotypes shared with exosome and TRAMP mutants (Paul and Montpetit, 2016). Enp1p is a ribosome biogenesis factor required for 40S pre-ribosome assembly and its export to the cytoplasm (Liang and Fournier, 1995; Dragon *et al.*, 2002; Chen *et al.*,

2003), while Srm1p functions in nucleocytoplasmic transport (Amberg *et al.*, 1993). Inactivation of Enp1p led to shared phenotypes with exosome mutants that include poly(A)-RNA, mRNA, and RBP (e.g. Nab2p) accumulation in the nucleolus. Similarly, loss of Srm1p caused poly(A)-RNA and mRNA accumulation in the nucleolus but differed from *enp1-1* mutants in that Nab2p is largely excluded from the nucleus due to disrupted nucleocytoplasmic transport. It may be that nuclear exclusion of RNA biogenesis factors, or sequestration of the same factors in a sub-nuclear poly(A)-RNA dense compartment, has the same impact on the cell, explaining the shared terminal phenotypes of exosome, TRAMP, *enp1-1*, and *srm1-ts* mutants (Paul and Montpetit, 2016). However, it is currently unclear how these shared phenotypes arise and if they represent a common cellular state at the molecular level.

In this work, analyses of *dis3-1*, *csl4-ph, rrp6Δ, enp1-1*, and *srm1-ts* yeast strains showed that these mutants have similar transcriptomes following growth at 37°C, including correlated changes in mRNA expression, levels of pervasive transcripts, and stable polyadenylation of ncRNAs. Inactivation of the exosome or Enp1p further led to Nab2p binding pre-ribosomal RNAs, snoRNAs, and ribosomal processing factors, with rapid inhibition of RNAPI transcription rescuing both the nucleolar poly(A)-RNA accumulation and Nab2p localization phenotypes. These observations together with localization of poly(A)-RNA and Nab2p in the nucleolus suggest that errors in rRNA biogenesis in a *csl4-ph* or *enp1-1* mutant lead to the generation of poly(A)-RNA species that engage Nab2p. In line with these observations, overexpression of another PABP, Pab1p, suppressed the nucleolar poly(A)-RNA accumulation phenotype and improved the growth of *enp1-1* strains at a semi-permissive temperature. These data provide evidence that stabilized nuclear poly(A)-RNA is toxic to the cell likely due to competition for essential RNA biogenesis factors. Moreover, the convergence on a shared set of phenotypes in mutants that function in relatively distinct biological process provides evidence that multiple initiating events with unique molecular origins can ultimately lead to a generalized failure in nuclear RNA homeostasis.

4.2 Results

4.2.1 enp1-1 and *srm1-ts* mutants have transcriptome profiles similar to exosome mutants

The disruption of multiple genes in *S. cerevisiae*, including *CSL4, DIS3*, *RRP6, ENP1,* and *SRM1,* which function in a set of distinct biological processes were identified to result in a common set of phenotypes that included poly(A)-RNA and mRNA accumulation in the nucleolus (Carneiro *et al.*, 2007; Paul and Montpetit, 2016). The impact of losing Dis3p function on the transcriptome has previously been described (Gudipati *et al.*, 2012), as has the deletion of *RRP6* at 30°C (Fox *et al.*, 2015; Vera and Dowell, 2016). These data have highlighted the diverse functions of the exosome in the processing and decay of ncRNAs (e.g. snoRNAs and rRNAs), mRNAs, and pervasive transcripts. However, transcriptomes of *csl4-ph, enp1-1, srm1-ts*, and *rrp6Δ* at 37°C, where these mutants share RNA export and RBP mis-localization phenotypes, have not been analyzed. Therefore, RNA-seq analyses were performed with ribo-depleted RNA samples prepared from control, *csl4-ph, dis3-1*, *rrp6Δ*, *enp1-1,* and *srm1-ts* strains after 90 minutes at 37°C. At this timepoint, poly(A)-RNA accumulation was apparent in all mutants (Figure 4-1A). Before proceeding to the analyses of RNA-seq data, the quality of biological replicates was checked by hierarchically clustering using the Z score of log transformed raw read counts from all transcripts, which showed that biological replicates clustered together, with replicates of control strains being the most disparate (Figure 4-1B). Further analysis of the RNA-seq data from ribosome depleted RNA samples showed that within the control strain ~87% of reads were generated from mRNAs and the remaining ~13% from ncRNAs, which included rRNAs, snRNAs, snoRNAs, and pervasive transcripts (Table 4-1 & Figure 4-2A). In mutants, the number of reads generated from mRNAs was generally lower, while ncRNA reads were expanded largely due to an increased level of pervasive transcripts.

To further investigate gene expression changes, transcript levels in mutants were compared to control to generate log2 fold change (log2FC) values for each transcript with an associated p-value. P-values < 0.01 were considered as significant, with results shown and discussed here only involving transcripts that meet this criterion. At the level of individual mRNAs and ncRNAs, all five mutants showed correlated changes in gene expression as compared to the control, including a 0.91 correlation coefficient for *dis3-1* vs. *csl4-ph,* and

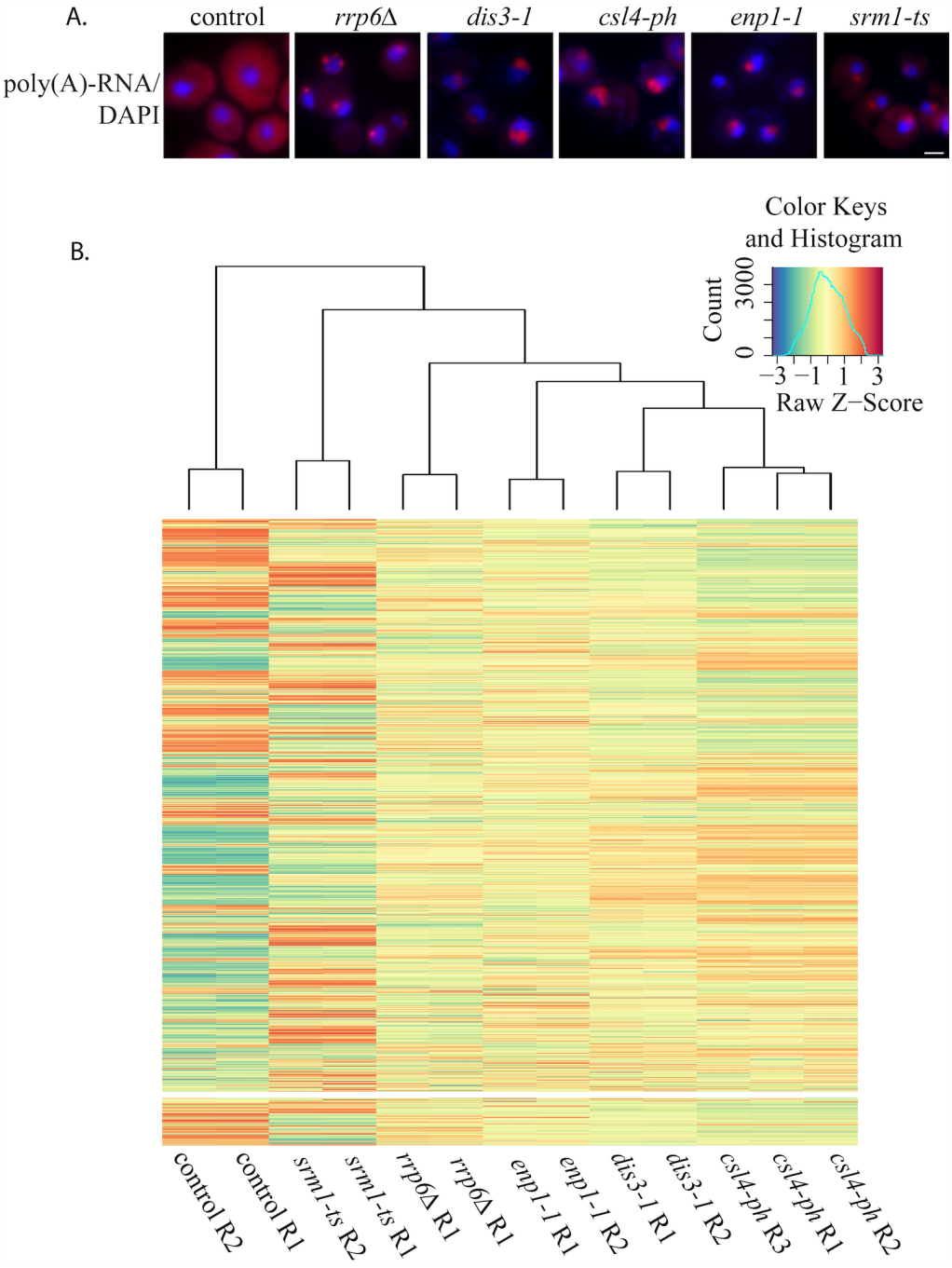


Figure 4-1. Poly(A)-RNA accumulation phenotypes and similarities among RNA-seq biological replicates. (A) FISH images showing nuclear accumulation of poly(A)-RNA in control, *csl4-ph*, *enp1-1*, *srm1-ts*, *rrp6Δ* and *dis3-1* strains after 90 minutes at 37°C. Nucleus shown in blue and poly(A)-RNA in red. Scale bar = 2 μm. (B) Heatmap of Z scores from log transformed read counts for all transcripts, which includes both mRNA and ncRNA. Columns of the heatmap correspond to sample (R=replicate) after shifting to 37°C for 90 minutes and the rows corresponds to transcripts.

Table 4-1: Percentage of reads from ribosome depleted RNA-seq data mapped to different classes of RNA

	control	*csl4-ph*	*enp1-1*	*srm1-ts*	*dis3-1*	*rrp6Δ*
pervasive transcripts	8% (1984569)	22% (6101041)	14.9% (6527177)	14% (5646349)	18% (3244726)	13% (3542114)
mRNA	87% (20424324)	72% (19477401)	82% (35942205)	79% (32870829)	76.9% (14048008)	82% (22737534)
rRNA	0.1% (40156)	1% (319473)	0.1% (111372)	1% (444812)	0.1% (24354)	0.1% (49818)
snoRNA	1.9% (524634)	3% (707868)	2% (815044)	3% (1306683)	3% (487912)	2.9% (885783)
snRNA	2% (445507)	2% (554309)	2% (665408)	3% (1425385)	2% (346079)	2% (596947)

correlations of 0.84 (*dis3-1* vs. *enp1-1*) and 0.85 (*csl4-ph* vs. *enp1-1*) between core exosome subunits and *enp1-1* (Figure 4-2B). K-means clustering (K=6) using log2FC data further showed that the same groups of transcripts were up- (cluster 1, 3-5) and down-regulated (cluster 2 & 6) in these mutants, with pervasive transcripts tending to increase, while mRNAs decreased (Figure 4-2C). Pervasive transcripts can be grouped into multiple classes (i.e. CUTs, SUTs, XUTs, and NUTs) based on the machineries their biogenesis and decay are dependent upon (Xu *et al.*, 2009; Van Dijk *et al.*, 2011; Schulz *et al.*, 2013). The generation of distribution graphs and log2FC correlation co-efficient heatmaps for each class showed that transcriptome changes were similar for each group of pervasive transcripts in the mutants tested (Figure 4-2D, Figure 4-3), suggesting that no one class of pervasive transcript was being specifically impacted by these mutants. Together, these data show that the phenotypic similarities between exosome mutants (i.e. *csl4-ph, dis3-1*, and *rrp6Δ*) and the ribosome biogenesis mutant *enp1-1,* and nucleocytoplasmic shuttling mutant *srm1-ts,* extend to include common changes in the cellular transcriptome.

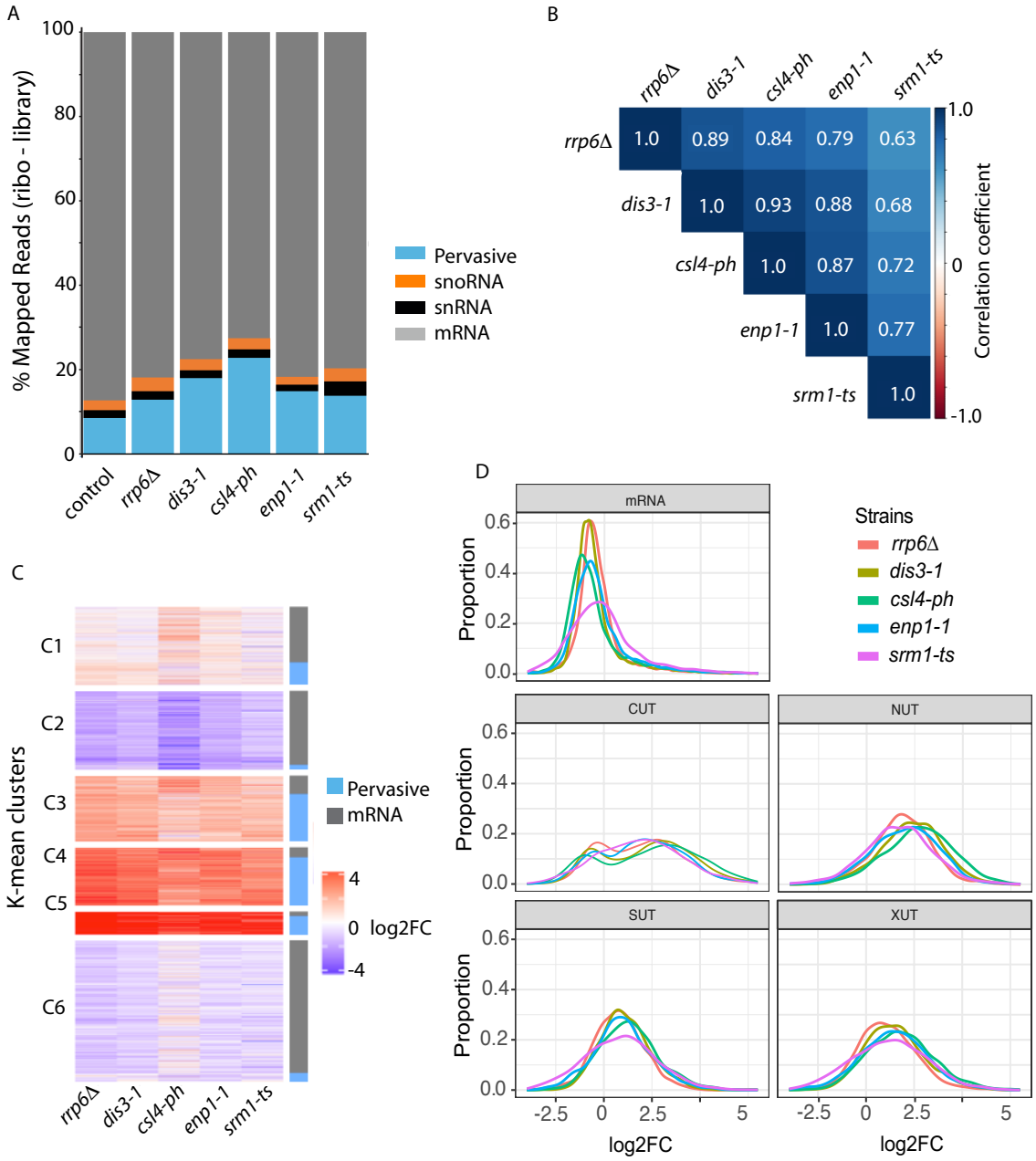


Figure 4-2. Similarities in the transcriptome of RNA decay and ncRNA processing mutants. (A) Percent of reads mapped to different classes of RNAs from ribosome depleted libraries of mutant strains (*csl4-ph*, *enp1-1*, *srm1-ts*, *rrp6Δ*, and *dis3-1*) after growth at 37°C for 90 minutes. (B) Heatmap showing the Pearson correlation co-efficient of log2FC values from comparing the control strain with the mutant strains. (C) Heatmap showing k-means clustering of log2FC in the mutant strains after 90 minutes of growth at 37°C with respect to the control strain. Within the heatmap, red indicates positive and purple indicates negative log2FC values. The side bar indicates pervasive transcripts (in blue) and mRNAs (in gray). (D) Distribution of log2FC values in the mutant strains for reads mapping to mRNAs and defined classes of pervasive transcripts (i.e. CUTs, SUTs, NUTs, and XUTs).

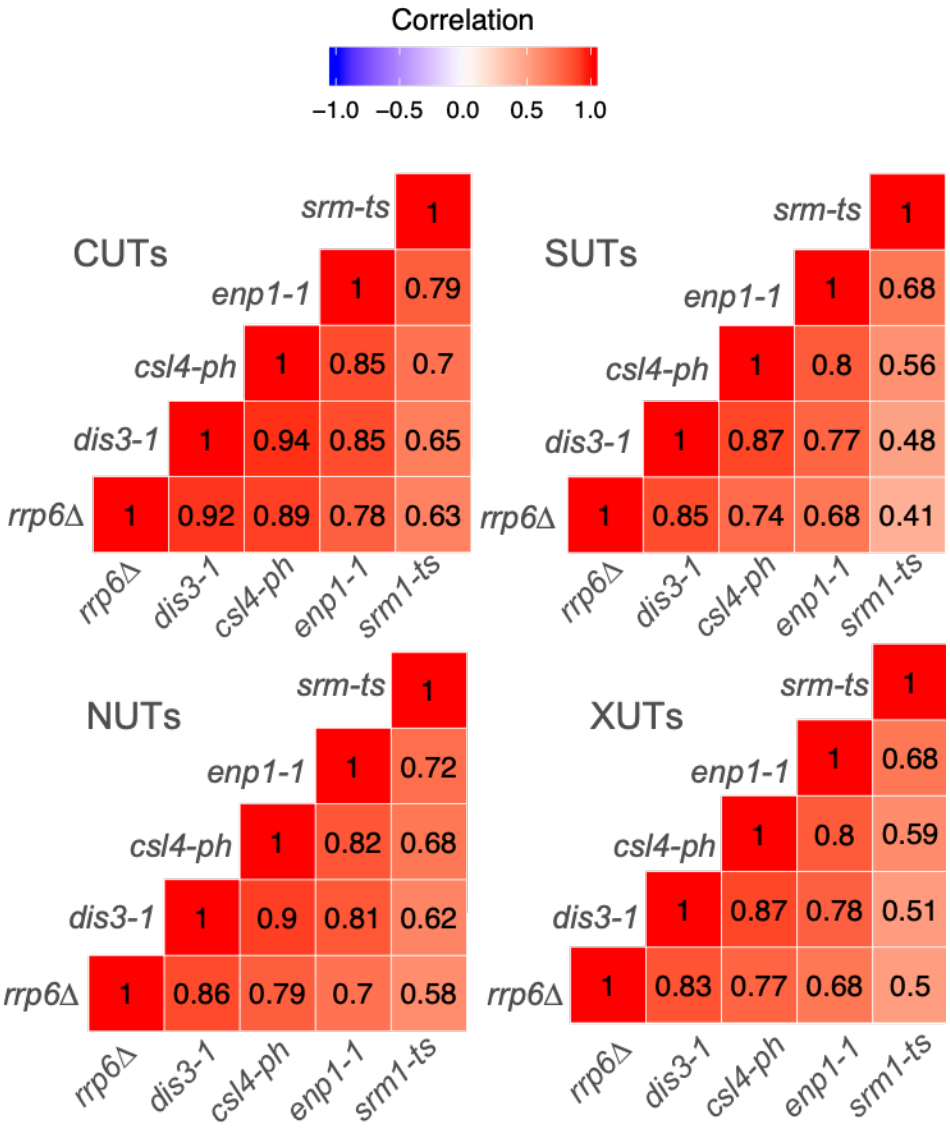


Figure 4-3. Similarities in pervasive transcript expression in RNA decay and ncRNA processing mutants. Heatmaps show the Pearson correlation co-efficient from log2FC values calculated by comparing the RNA-seq data from control strain and the mutant strains after 90 minutes of growth at 37°C for four classes of pervasive transcripts, CUTs (top left), SUTs (top right), NUTs (bottom left), and XUTs (bottom right)

4.2.2 Changes in mRNA abundance reflect RBP-binding profiles

mRNA expression was generally lower in *csl4-ph, dis3-1*, *rrp6Δ, enp1-1,* and *srm1-ts* strains when shifted to the non-permissive temperature (Figure 4-2). Recent work has defined ten broad classes of mRNA based on RBP interaction profiles (Tuck and Tollervey, 2013), with each class likely being indicative of differences in mRNA processing. For example, class I-III mRNAs tend to be of low abundance in cells and bind specific RBPs (e.g. the TRAMP component Mtr4p) in a pattern that is similar to pervasive transcripts that are targeted by the exosome for degradation (Tuck and Tollervey, 2013). In contrast, class X tends to be highly expressed and bind cytoplasmic RBPs, suggestive of transcripts that are exported, translated, and decayed in the cytoplasm (Tuck and Tollervey, 2013). Consequently, to understand if gene expression changes correlate with RBP interaction profiles, mRNA expression changes were compared within each mRNA class for the five *Ts* mutants (Figure 4-4A). From this analysis it was shown that mRNAs found in classes I-III were generally expressed at higher levels, while mRNAs in class IV-X trended lower in expression (Figure 4-4A). Based on gene ontology analyses, stabilized mRNAs in class I included meiosis specific transcripts, which are known to be targeted by the exosome in vegetative cells (Lardenois *et al.*, 2011). The stabilization of class I-III mRNAs is suggestive of a loss of nuclear surveillance activity in these five mutants, which results in the stabilization of mRNA and pervasive transcripts that are normally decayed and kept at low to non-detectable levels. In the case of down-regulated mRNA, by plotting mRNAs using relative expression levels (RPKM) in the control strain (x-axis) to the fold change seen in a mutant (y-axis), it is also observed that the most highly expressed mRNAs (class IX-X) are strongly downregulated (Figure 4-4B). This observed impact on a large number of genes that are highly expressed suggests that down regulation of these transcripts may broadly result from reduced levels of mRNA biogenesis. This may be due in part to changes in the availability of RBPs required to support transcription and mRNA processing through sequestering RBPs in the nucleolar compartment, as reported for Nab2p and Prp19p in *dis3-1* and *enp1-1* strains (Paul and Montpetit, 2016). Down regulation of transcription may also represent activation of specific stress pathways to protect the cell in response to disrupted nuclear RNA processing, which was recently reported in ribosome processing mutants (Albert *et al.*, 2019).

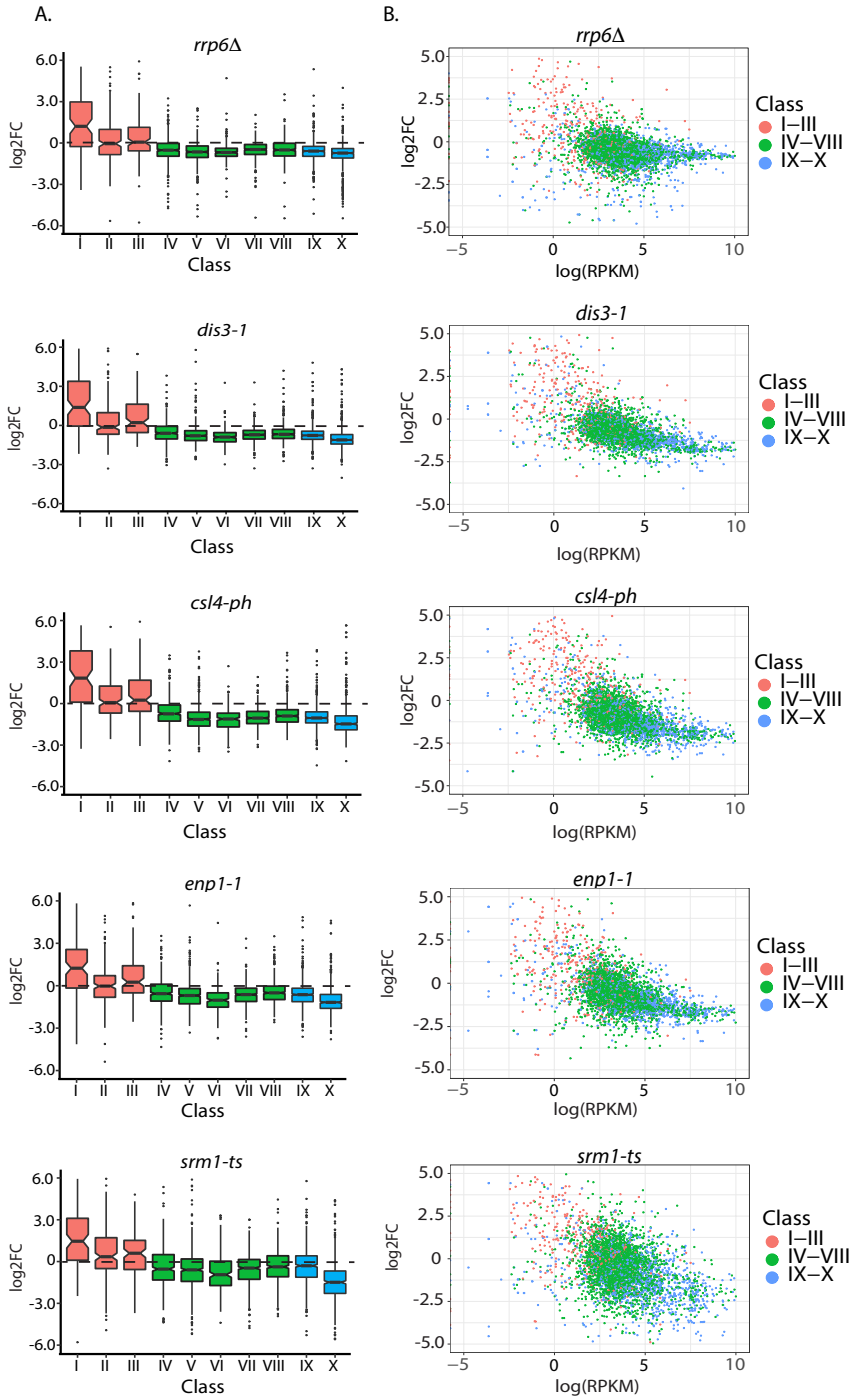


Figure 4-4. Variations in transcript abundance in RNA decay and ncRNA processing mutants based on mRNP class. (A) Boxplot representing log2FC values (y-axis) in the mutant strains *rrp6Δ*, *dis3-1*, *srm1-ts*, *csl4-ph*, and *enp1-1* (from top to bottom) for different classes of mRNA (x-axis) defined by RBP interaction patterns (Tuck and Tollervey, 2013). (B) Scatter plot showing the relationship between log2FC values (y-axis) in a mutant strain *vs.* mRNA abundance (x-axis) in the control strain. Color of the points represent different classes of mRNA transcripts, orange (I-III), green (IV-VIII), and blue (IX-X). RNA abundance is displayed as a log RPKM value, which is a read count normalized with respect to the total number of reads and the length of a gene.

4.2.3 Csl4p is central to the exosome protein interaction network

The correlated changes in mRNA and ncRNA expression between *dis3-1* and *csl4*-ph suggest that loss of Csl4p activity broadly impacts exosome mediated RNA processing and decay in a manner very similar to that resulting from loss of Dis3p activity. As such, a general characterization of the *csl4-ph* allele was undertaken to understand how exosome function may be impacted in this mutant. Sequencing revealed mutations in the *csl4-ph* allele that resulted in amino acid changes T23A, R206S, F224S, D251G, P278R, and A282P. Based on published structural models of Csl4p and the exosome (Makino *et al.*, 2013; Wasmuth *et al.*, 2014; Kowalinski *et al.*, 2016), these mutations are found within the core fold of Csl4p and at various contact points between Csl4p and the exosome complex (data not shown), which may result in disruptions to the overall structure of Csl4p and binding of Csl4p to the exosome complex. In support of this, fluorescence imaging to detect Csl4p-GFP showed that the mutant allele was expressed and was properly localized at 25°C, but at 37°C, Csl4p was rapidly degraded becoming undetectable after 30-60 minutes at the non-permissive temperature (Figure 4-5).

To understand what impact loss of Csl4p has on the integrity of the exosome, affinity purification-mass spectrometry (AP-MS) was employed using a tagged exosome component (Rrp41p-Protein A) to investigate both the assembly state of the exosome and interactions occurring with binding partners. Following growth at 37°C for 90 minutes, Rrp41p isolated from *csl4-ph* cells showed a reduced protein interaction network as compared to control (Figure 4-6), with only core exosome proteins (Dis3p, Rrp43p, Rrp4p, Rrp40p, Rrp42p, Rrp46p, Mtr3p, Rrp45p) remaining associated with Rrp41p. Specific Rrp41p binding partners lost in *csl4-ph* included Rrp6p, the TRAMP complex component Mtr4p, the exosome associated factor Mpp6p, the exosome adaptor Utp18p, and numerous ribosomal maturation factors involved in pre-rRNA processing. These results suggest that loss of Csl4p causes the majority of the direct and indirect protein interactions occurring between the exosome and other RNA processing machineries to be destabilized or lost. This finding is supported by structural data showing a role for Csl4p in anchoring Rrp6p and Mtr4p to the exosome (Makino *et al.*, 2013; Thoms *et al.*, 2015; Kowalinski *et al.*, 2016; Zinder *et al.*, 2016). Given that Csl4p directly engages some RNA substrates (Schneider *et al.*, 2012), Mtr4p binds substrates that are processed by both the threaded and direct-access pathways (Thoms *et al.*, 2015; Falk *et al.*, 2017b; Weick *et al.*, 2018), and Mtr4p

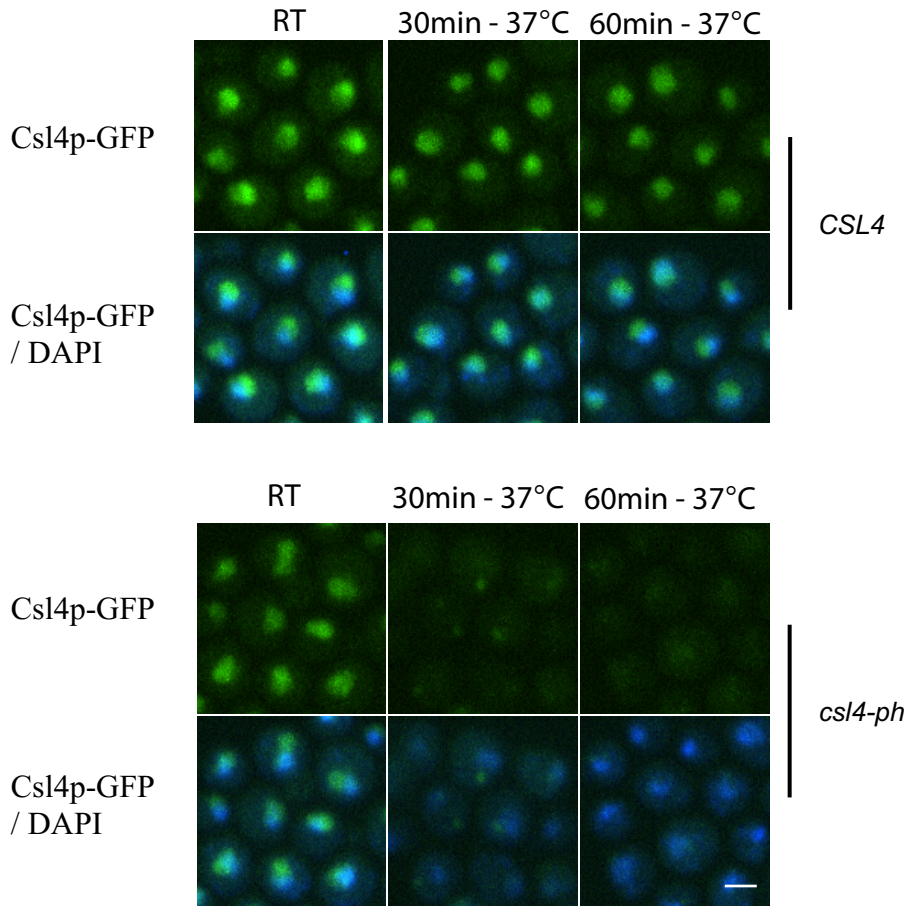


Figure 4-5. Csl4p-GFP localization in control and *csl4-ph* strains. Localization of Csl4p-GFP in the control strain (top) versus the *csl4-ph* strain (bottom) at room temperature and after 30 or 60 minutes of growth at 37°C. The nucleus is shown in blue and Csl4p-GFP in green. Scale bar = 2 μm.

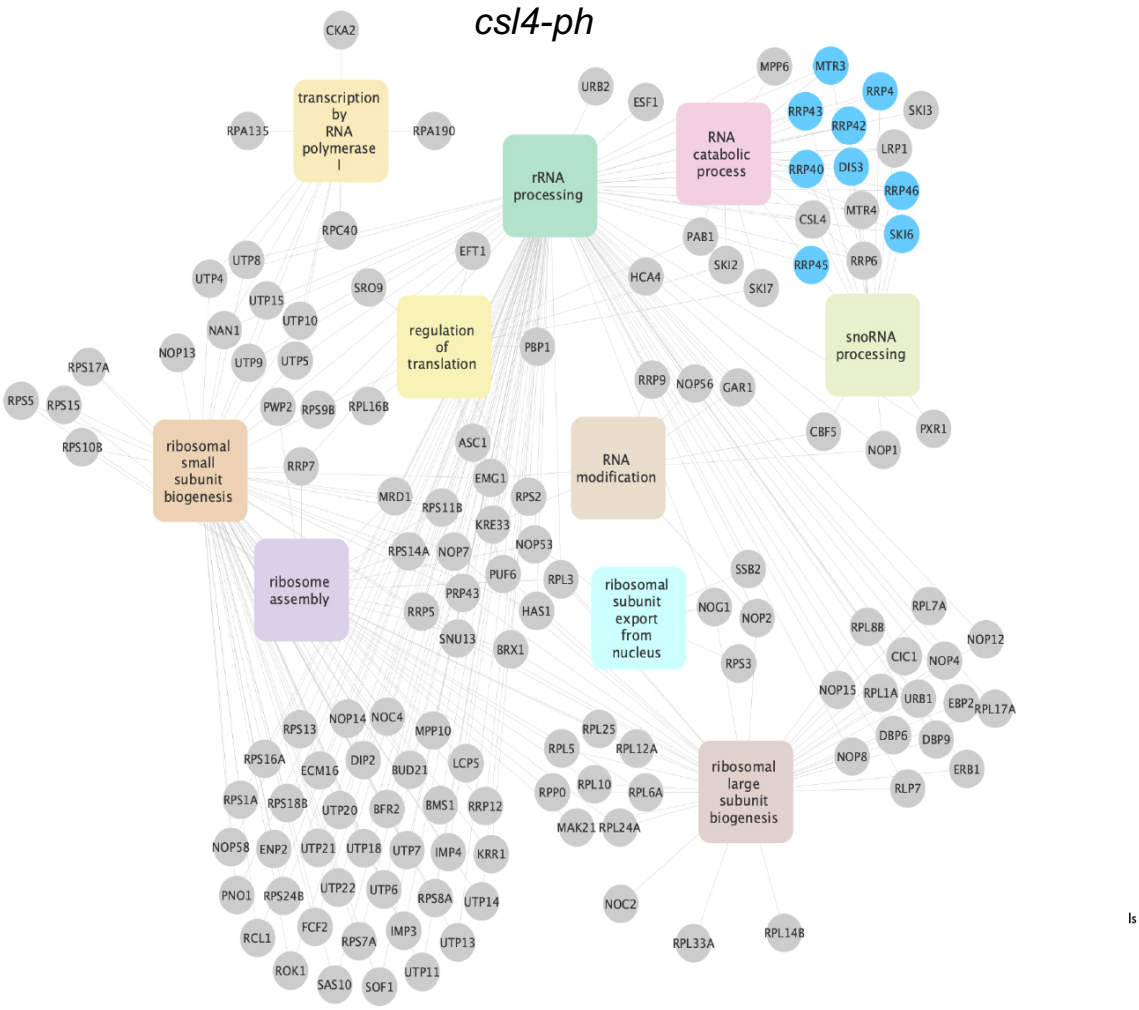


Figure 4-6. Rrp41p protein-protein interactions in control and *csl4-ph* strains. Protein interaction network displaying proteins (rounds nodes) identified by AP-MS in co-immunoprecipitation experiments with Rrp41p-Protein A. Nodes in gray represent proteins found with Rrp41p in wild-type cells, while blue nodes were those proteins identified in both wild-type and *csl4-ph* strains grown at 37°C for 90 minutes. Proteins were manually organized around gene ontology (GO) categories to aid data interpretation (GO terms shown in rectangles). AP-MS was performed by Carolina Aguilar (IRCM).

anchors adaptor proteins to the exosome for rRNA processing (Houseley and Tollervey, 2006; Klauer and van Hoof, 2013; Thoms *et al.*, 2015), it is likely that loss of Csl4p abolishes exosome activity on a large fraction of RNA substrates. This is in agreement with the highly correlated changes in the cellular transcriptome of *csl4-ph* and *dis3-1* strains involving mRNAs, pervasive transcripts, and ncRNAs (Figure 4-2 and 4-3).

In comparison, the Rrp41p-Protein A interaction network in *dis3-1* and *enp1-1* strains were largely intact, while *mtr4-1* and *srm1-ts* strains showed a more severe reduction in observed interactions (Figures 4-7, 4-8, 4-9, and 4-10). For *srm1-ts*, loss of the majority of the interactions may be a result of failed nucleocytoplasmic transport and an inability of the exosome to access the nucleus. To address this, the localization of Rrp41p-GFP was assessed in *srm1-ts* strain at the permissive and non-permissive growth temperature, which indeed showed that nuclear Rrp41p-GFP levels were largely reduced at 37°C in this mutant (Figure 4-11). The loss of Rrp41p-GFP, and potentially the exosome as a whole, could explain the exosome-like phenotypes observed in the *srm1-ts* mutant.

4.2.4 Accumulation of polyadenylated ncRNAs

The shared phenotypes and transcriptomic profiles of *csl4-ph, rrp6Δ, dis3-1, enp1-1, and srm1-ts* mutants suggest that the nucleolar poly(A)-RNA signals in these mutants may arise from stabilized pervasive transcripts and mRNAs normally decayed during vegetative growth. However, it remains unknown which transcripts are specifically polyadenylated in these mutants and are responsible for the nucleolar poly(A)-RNA signal that is observed by FISH at non-permissive temperature (Figure 4-1A). To characterize polyadenylated transcripts, RNA-seq was performed using dT-purified RNA from control, *csl4-ph, enp1-1,* and *srm1-ts* strains (Figure 4-12A & Table 4-2). In control cells, dT selection resulted in sequencing data with ~89% of reads being mapped to mRNAs and ~5% mapping to pervasive transcripts (i.e. CUTs, NUTs, SUTs and XUTs) (Marquardt *et al.*, 2011; Van Dijk *et al.*, 2011). The remaining ~5% reads mapped to rRNAs, indicative of polyadenylated ribosomal RNA precursors generated during rRNA biogenesis (Tudek *et al.*, 2018a). In the case of *csl4-ph* and *enp1-1*, the percent of reads mapped to pervasive transcripts increased 2.6- and 2.2-fold in comparison to control strains, as was

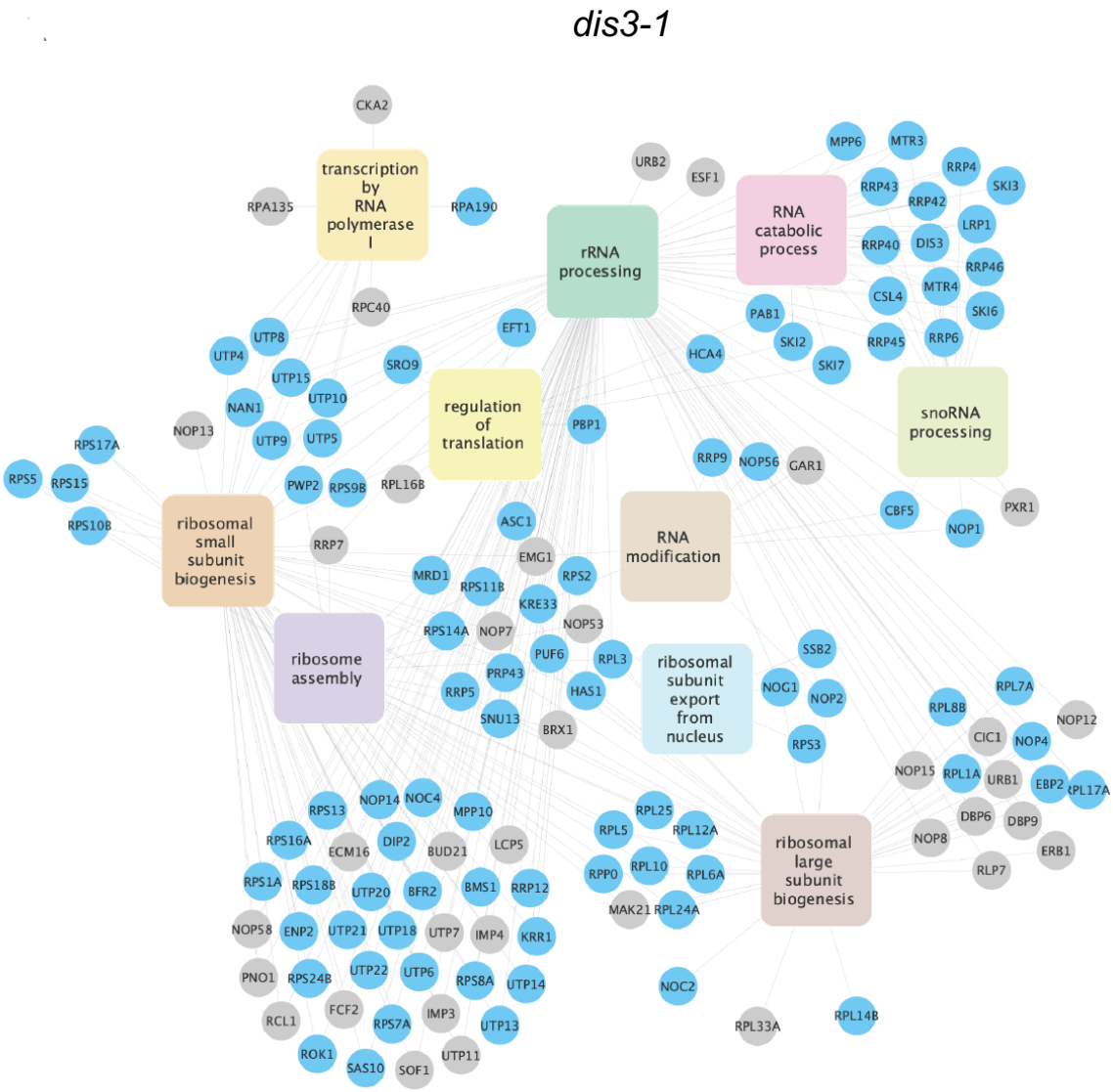


Figure 4-7. Rrp41p protein-protein interactions in control and *dis3-1* strains. Protein interaction network displaying proteins (rounds nodes) identified by AP-MS in co-immunoprecipitation experiments with Rrp41p-Protein A. Nodes in gray represent proteins found with Rrp41p in wild-type cells, while blue nodes were those proteins identified in both wild-type and *dis3-1* strains grown at 37°C for 90 minutes. Proteins were manually organized around gene ontology (GO) categories to aid data interpretation (GO terms shown in rectangles). AP-MS was performed by Carolina Aguilar (IRCM).

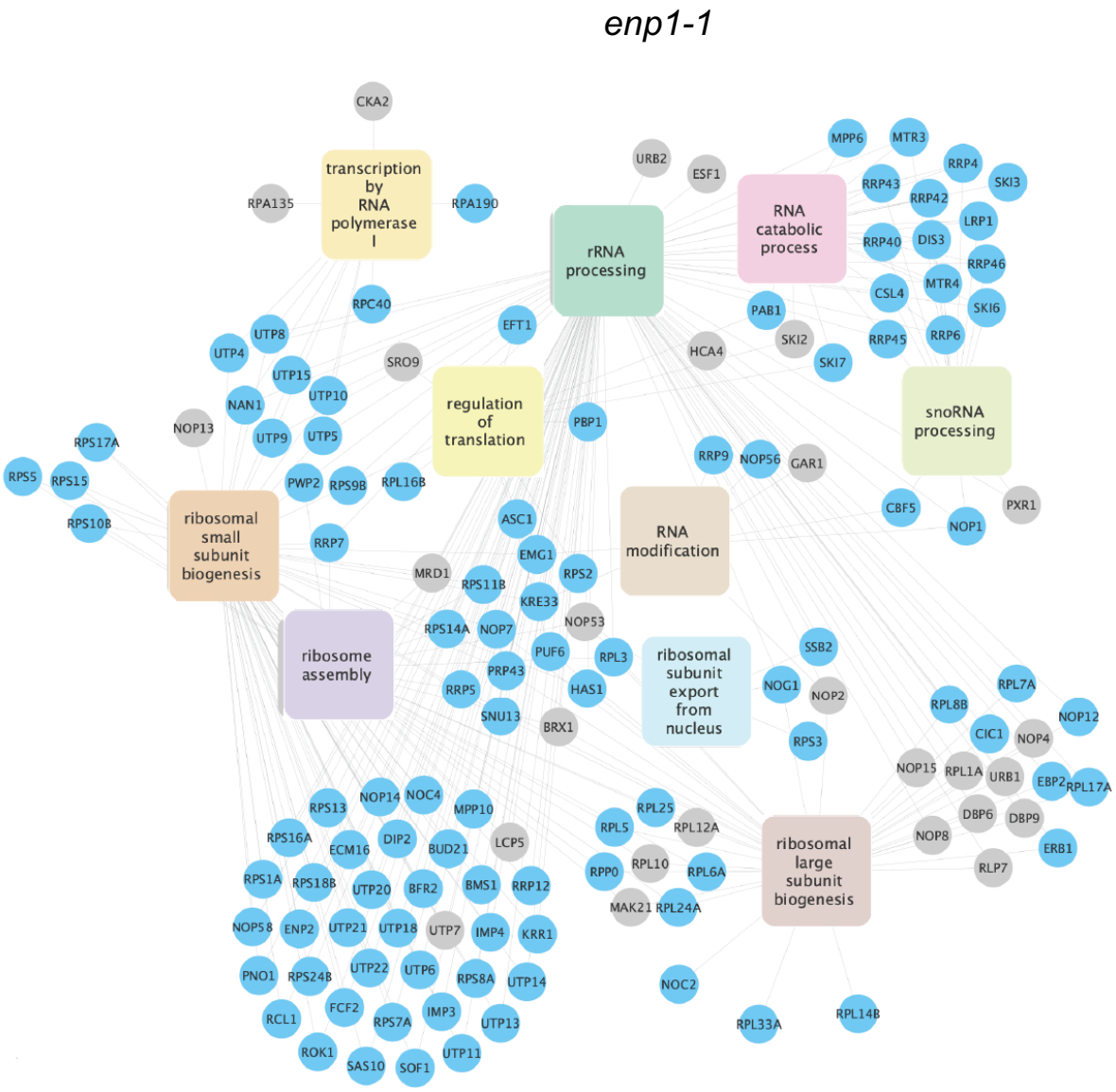


Figure 4-8. Rrp41p protein-protein interactions in control and *enp1-1* strains. Protein interaction network displaying proteins (rounds nodes) identified by AP-MS in co-immunoprecipitation experiments with Rrp41p-Protein A. Nodes in gray represent proteins found with Rrp41p in wild-type cells, while blue nodes were those proteins identified in both wild-type and *enp1-1* strains grown at 37°C for 90 minutes. Proteins were manually organized around gene ontology (GO) categories to aid data interpretation (GO terms shown in rectangles). AP-MS was performed by Carolina Aguilar (IRCM).

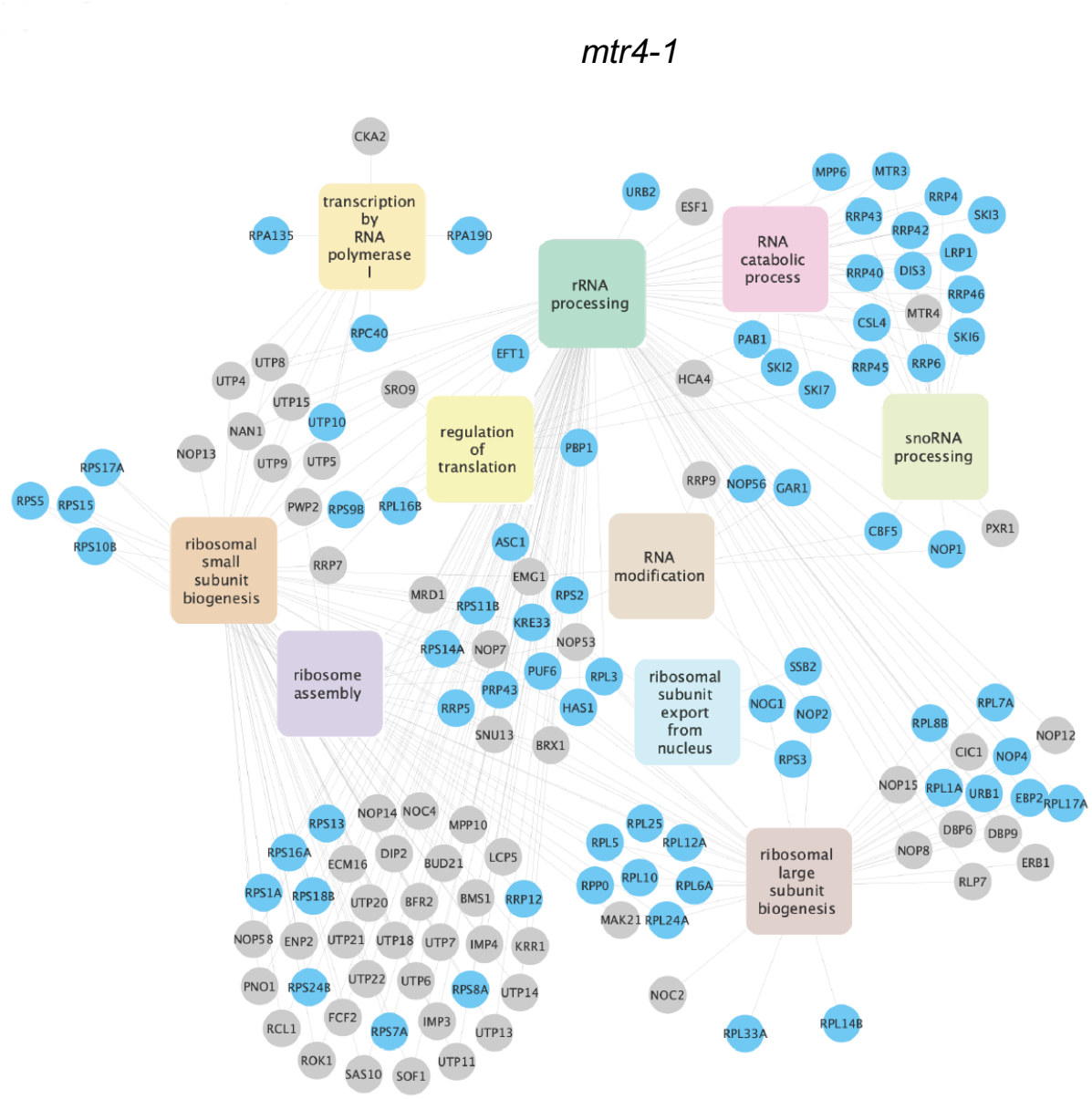


Figure 4-9. Rrp41p protein-protein interactions in control and *mtr4-1* strains. Protein interaction network displaying proteins (rounds nodes) identified by AP-MS in co-immunoprecipitation experiments with Rrp41p-Protein A. Nodes in gray represent proteins found with Rrp41p in wild-type cells, while blue nodes were those proteins identified in both wild-type and *mtr4-1* strains grown at 37°C for 90 minutes. Proteins were manually organized around gene ontology (GO) categories to aid data interpretation (GO terms shown in rectangles). AP-MS was performed by Carolina Aguilar (IRCM).

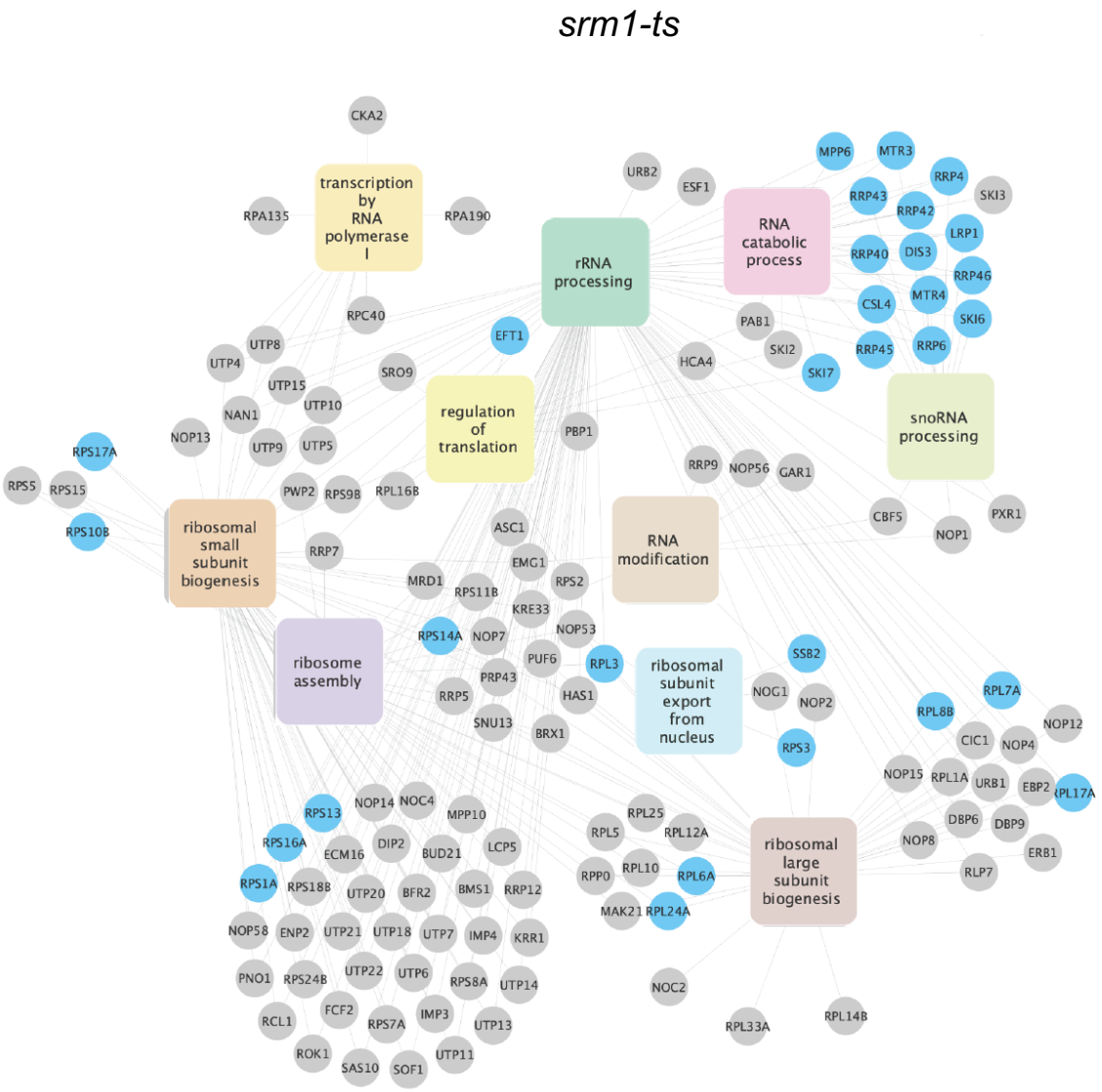


Figure 4-10. Rrp41p protein-protein interactions in control and *srm1-ts* strains. Protein interaction network displaying proteins (rounds nodes) identified by AP-MS in co-immunoprecipitation experiments with Rrp41p-Protein A. Nodes in gray represent proteins found with Rrp41p in wild-type cells, while blue nodes were those proteins identified in both wild-type and *srm1-ts* strains grown at 37°C for 90 minutes. Proteins were manually organized around gene ontology (GO) categories to aid data interpretation (GO terms shown in rectangles). AP-MS was performed by Carolina Aguilar (IRCM).

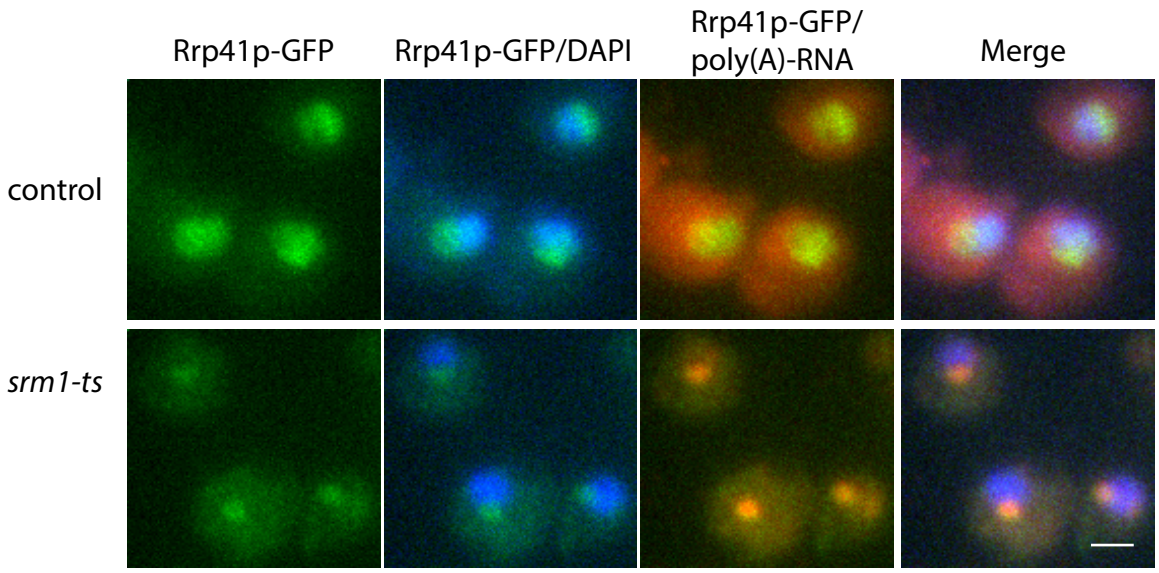


Figure 4-11. Rrp41p-GFP localization in *srm1-ts*. Localization of Rrp41p-GFP (green) and poly(A)-RNA (red) in reference to DNA (blue) in the control and *srm1-ts* strains grown at 37°C for 90 minutes. Scale bar = 2μm.

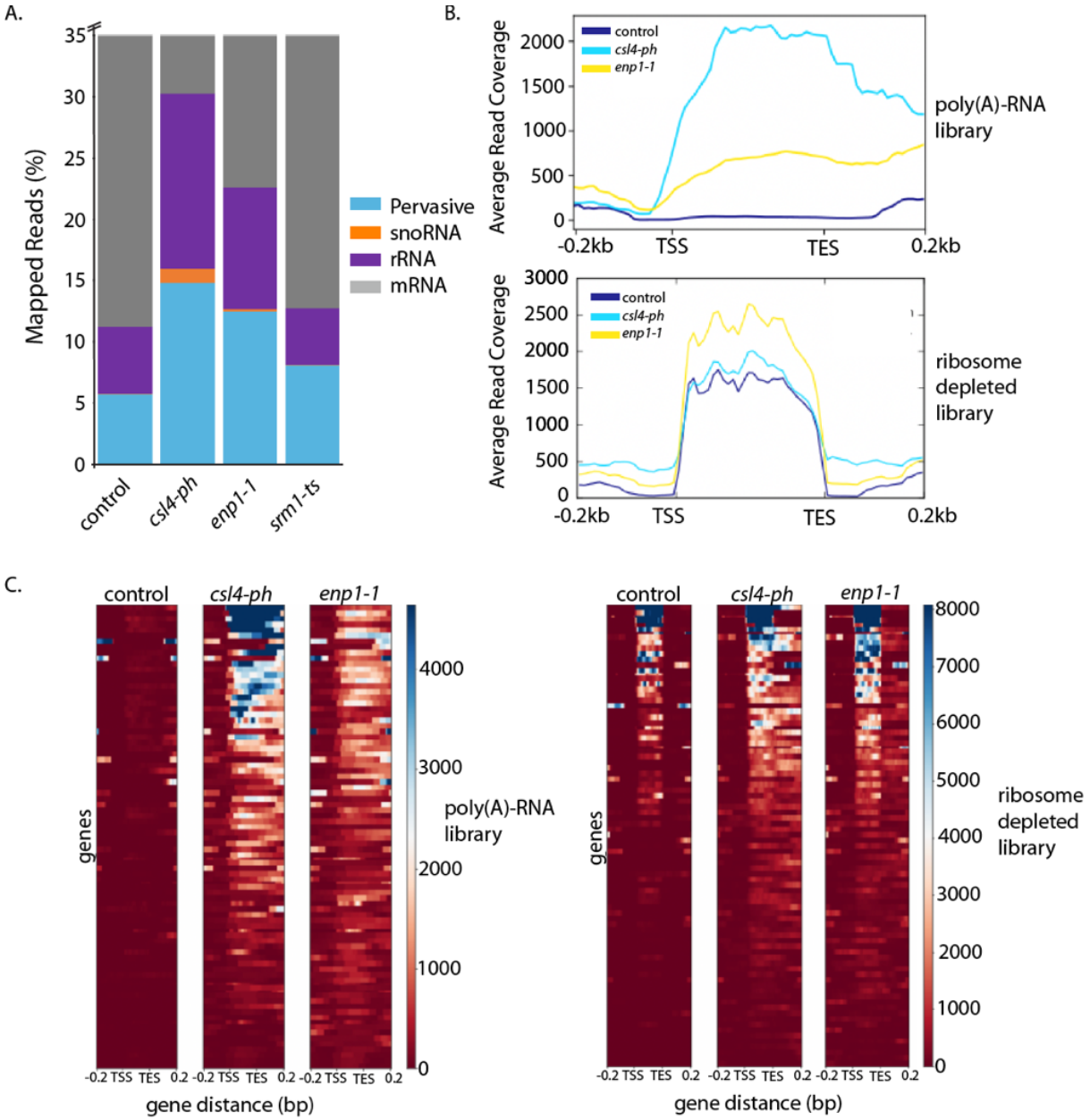


Figure 4-12. Analysis of polyadenylated transcripts from dT purified RNA-seq libraries in the *csl4-ph, enp1-1,* and *srm1-ts* mutant strains. (A) Percentage of reads from poly(A) selected RNA-seq libraries in different RNA classes in *control*, *csl4-ph*, *enp1-1,* and *srm1-ts* strains grown at 37°C for 90 minutes. (B) Metagene plot showing average read density across snoRNA genes and 200 nucleotides upstream and downstream of the transcription start site (TSS) and transcription end site (TES). Data for control, *csl4-ph*, and *enp1-1* strains grown at 37°C for 90 minutes both in ribosome depleted (bottom) and poly(A) selected (top) RNA-seq libraries are shown. (C) Heatmap showing read density across all snoRNA genes and 200 nucleotides upstream and downstream of TSS and TES in poly(A) selected (left) and ribosome depleted (right) RNA-seq libraries in control, *csl4-ph*, and *enp1-1* strains grown at 37°C for 90 minutes.

Table 4-2. Percentage of reads from poly(A) selected RNA-seq libraries mapping to different classes of RNA.

	control	*csl4-ph*	*enp1-1*	*srm1-ts*
mRNA	88.76% (22451629)	69.69% (19652291)	77.37% (32490208)	87.22% (41384907)
snRNA	0.006% (1576)	0.026% (7523)	0.025% (10815)	0.007% (3688)
snoRNA	0.04% (10124)	1.15% (325522)	0.175% (73606)	0.025% (12055)
rRNA	5.4% (1381999)	14.3% (4039376)	9.92% (4168065)	4.6% (2196483)
pervasive transcripts	5.72% (1447288)	14.8% (4174629)	12.5% (5247670)	8.11% (3849917)

observed in ribo-depleted samples (Figure 4-2A). The proportion of ncRNA reads increased to 30.3% and 22.6% of total reads in *csl4-ph* and *enp1-1* strains, including a 2.6- and 1.8-fold increase in the percentage of reads mapping to rRNA and 28.7 and 4.3-fold increase in percentage of snoRNA reads, changes that were not seen in data generated from ribosome depleted RNA samples (Table 4-1 and 4-2). The differences observed between these RNA-seq data sets suggesting that a subset of snoRNAs are polyadenylated in *csl4-ph* and *enp1-1* strains

To further compare data from these two types of RNA-seq libraries with respect to snoRNAs, average read counts were mapped across all snoRNAs from ribosome depleted vs. poly(A)-RNA enriched sequencing data. The resulting metagene plots showed that read numbers dramatically increased in poly(A)-RNA data from both *enp1-1* and *csl4-ph* strains within the gene body and downstream of the transcription end site (TES), while ribo-depleted data indicated that global snoRNA levels are not changed to a large degree (Figure 4-12B). Moreover, read density data at the level of individual snoRNAs in control, *csl4-ph,* and *enp1-1* libraries also indicated a difference between snoRNAs in these two types of RNA-seq libraries (Figure 4-12C). These observations show that steady-state snoRNA levels do not significantly change in the *csl4-ph* and *enp1-1* strains, based on ribosome depleted RNA-seq libraries, rather a proportion of snoRNAs in these mutants have processing defects that result in 3′-extended polyadenylated transcripts. Given the known role of the exosome in snoRNA processing (van Hoof *et al.*, 2000a; LaCava *et al.*, 2005; Kim *et al.*, 2006; Vasiljeva and Buratowski, 2006; Carneiro *et al.*, 2007;

Grzechnik and Kufel, 2008; Lemay *et al.*, 2010; Heo *et al.*, 2013), this may have been expected for *csl4-ph*; however, it is unexpected for *enp1-1*, which so far has only been implicated in rRNA biogenesis. In the case of *srm1-ts*, no significant changes in rRNA and snoRNA reads were observed and pervasive transcripts increased only to a small degree, suggesting that at the level of transcript polyadenylation *srm1-ts* is different from *csl4-ph* and *enp1-1* mutants. However, with the similarities in the ribosome depleted RNA-seq data between all these mutants, the weaker poly(A)-RNA signal observed by FISH in *srm1-ts* (Figure 4-1A), and the role of Srm1p in nucleocytoplasmic transport, it is possible that these results reflect failed nucleocytoplasmic transport. Specifically, failed import of proteins involved in performing (i.e. TRAMP complex) and directing (i.e. NNS complex) polyadenylation from accessing the nucleus in a *srm1-ts* mutant. To asses this, the localization of the TRAMP subunit Trf4p-GFP and NNS complex subunit Nrd1p-GFP were determined in *srm1-ts* cells, which showed that both proteins had a large cytoplasmic pool in this mutant at 37°C (Figure 4-13). It is expected that these differences in the localization of nuclear RNA processing factors are ultimately responsible for the both the observed RNA processing defects in *srm1-ts*, as well as the variance in ncRNA transcript polyadenylation between *srm1-ts* and the other mutants.

4.2.5 Nab2p engages pre-rRNAs and associated biogenesis factors

Changes in ncRNAs levels in poly(A) enriched RNA-seq data from *csl4-ph* and *enp1-1,* plus the localization of the poly(A)-RNA signal within the nucleolus, suggest that both may be a consequence of polyadenylated pre-rRNA and snoRNA transcripts accumulating in these mutants. To further demonstrate ncRNA polyadenylation, total RNA vs. oligo-dT purified RNA was analyzed by northern blotting. These data showed that *dis3-1*, *csl4-ph*, *rrp6Δ,* and *enp1-1* strains accumulated polyadenylated rRNAs (e.g. 23S) and snoRNAs (e.g. U14 and snR30), as compared to the *ACT1* mRNA (Figure 4-14). Specifically, in *enp1-1* cells, polyadenylated forms of late pre-40S rRNAs (17S, 21S and 20S; Figure 4-14, lanes 5 versus 12) were found to accumulate, in accordance with its role in late 40S subunit maturation (Choque *et al.*, 2018). Similarly, polyadenylated 40S precursors were increased in *csl4-ph* and *dis3-1* cells (Figure 4-14,

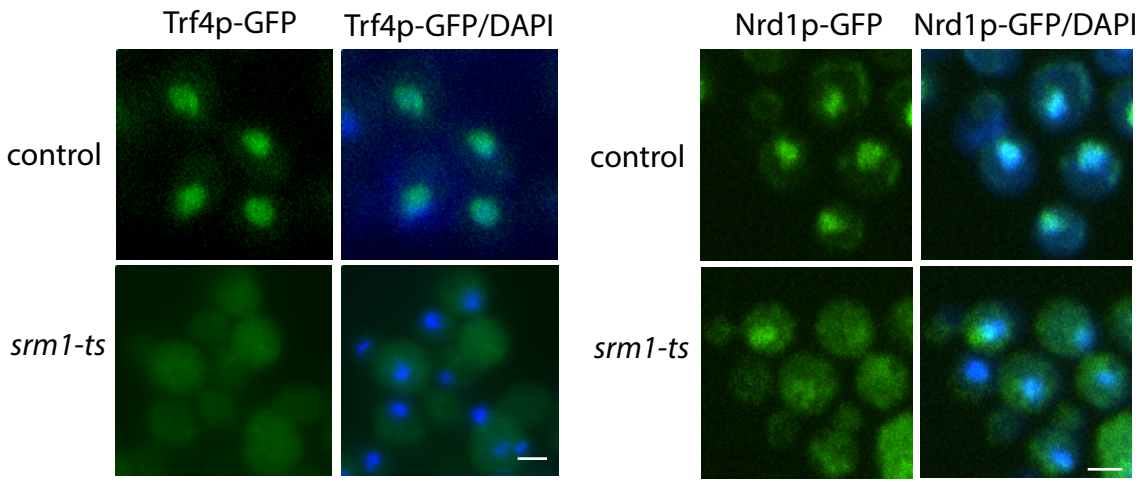


Figure 4-13. Localization of Trf4p and Nrd1p in *srm1-ts*. Localization of Trf4p-GFP (left) and Nrd1p-GFP (right) in control (top) and *srm1-ts* (bottom) strains with DAPI (blue) after 90 minutes temperature shift to 37 C. Scale bar = 2 μm.

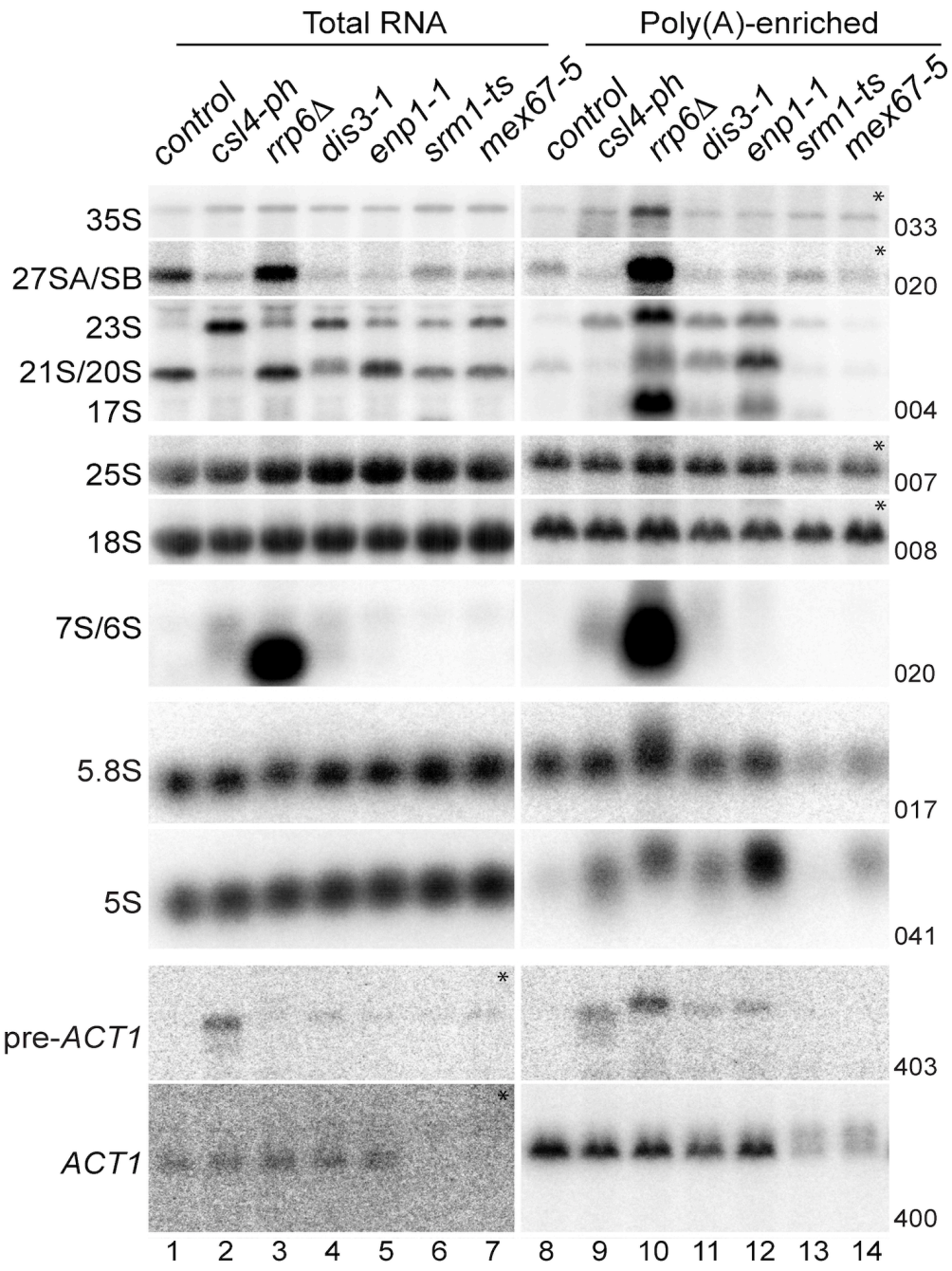


Figure 4-14. Polyadenylation of rRNAs. Northern blot showing pre-rRNAs (35S, 27SA/SB, 23S, 21S/20S, 17S, 7S/6S), *ACT1* mRNA, and mature rRNAs (18S, 25S, 5.8S, 5S) detected in control, *csl4-ph*, *rrp6Δ, dis3-1*, *enp1-1*, *srm1-ts,* and *mex67-5* strains after incubation at 37°C for 90 minutes from. Images are from blots loaded with total RNA (left) and poly(A) selected RNA (right). Numbers on the right side of the figure denote probes used to detect the targeted transcripts, which are detailed in Table 2.4. Lane numbers are given at the bottom of the figure. Northern blotting was performed by Carolina Aguilar (IRCM).

lanes 2 versus 9 and 4 versus 11), and both 40S and 60S precursors in *rrp6Δ* cells, (Figure 4-14, lanes 3 versus 10) pointing towards an inability to decay aberrant precursors due to the absence of exosome components (Zanchin and Goldfarb, 1999; Allmang *et al.*, 2000). No increase in polyadenylated ncRNAs was observed in *srm1-ts* (Figure 4-14, lanes 6 versus 13) or a mutant of mRNP export, *mex67-5* (Figure 4-14, lanes 7 versus 14). Northern blot of total and poly(A)-selected RNA were also used to test polyadenylation of snoRNA. Both U14 (C/D box snoRNA) and snR30 (H/ACA box snoRNA) showed enrichment in poly(A) selected total RNA in exosome and *enp1-1* (Figure 4-15A). In the case of snR30, an extension poly(A) test, referred to as an ePAT assay (Janicke *et al.*, 2012), was also used to detect the presence of poly(A) tails on the snR30, which is enriched ~5-fold in the dT selected RNA-seq data when compared to control. The ePAT assay also showed that snR30 was strongly polyadenylated in *csl4-ph* and *enp1-1* strains (Figure 4-15B). These data, together with the transcriptomics analyses, suggest that pre-ribosomal RNAs and snoRNAs are polyadenylated in both *csl4-ph* and *enp1-1* strains.

The increased levels of polyadenylated snoRNAs and pre-rRNAs (Figure 4-14, 4-15), accompanied by the re-localization of RBPs (e.g. Nab2p) to the nucleolus (Paul and Montpetit, 2016), suggests that this polyadenylated material may be bound inappropriately by RBPs normally associated with mRNA. To test for interactions between Nab2p and ncRNA transcripts in *csl4-ph* and *enp1-1* cells, northern analyses were performed on RNA purified with Nab2p following a shift to 37°C for 90 minutes. Using probes against rRNAs, a clear enrichment was seen for pre-rRNAs associated with Nab2p in *csl4-ph* and *enp1-1* cells compared to control, in particular 23S, 21S, 20S, and 5S (Figure 4-16). Similarly, snoRNAs U14 and snR30 associated with Nab2p in these mutant backgrounds (Figure 4-17). These interactions confirm that polyadenylated ncRNAs bind Nab2p, which could act to limit Nab2p activity in other pathways.

To further understand which RNPs Nab2p is associated with, AP-MS was performed with Nab2p-Protein A in control, *csl4-ph,* and *enp1-1* strains grown at 37°C for 90 minutes. As expected, MS data revealed that Nab2p interacted with proteins involved in mRNA transport and gene expression, including proteins of the nuclear pore complex and mRNA export pathway, in control cells (Figure 4-18). Nab2p isolated from *csl4-ph* maintained many of the same interactions but had an expanded interaction network that included proteins involved in rRNA processing and RNA surveillance, including the NNS and TRAMP complexes (Figure 4-18 and 4-19). In the case of *enp1-1,* the association of Nab2p with proteins of the nuclear pore complex

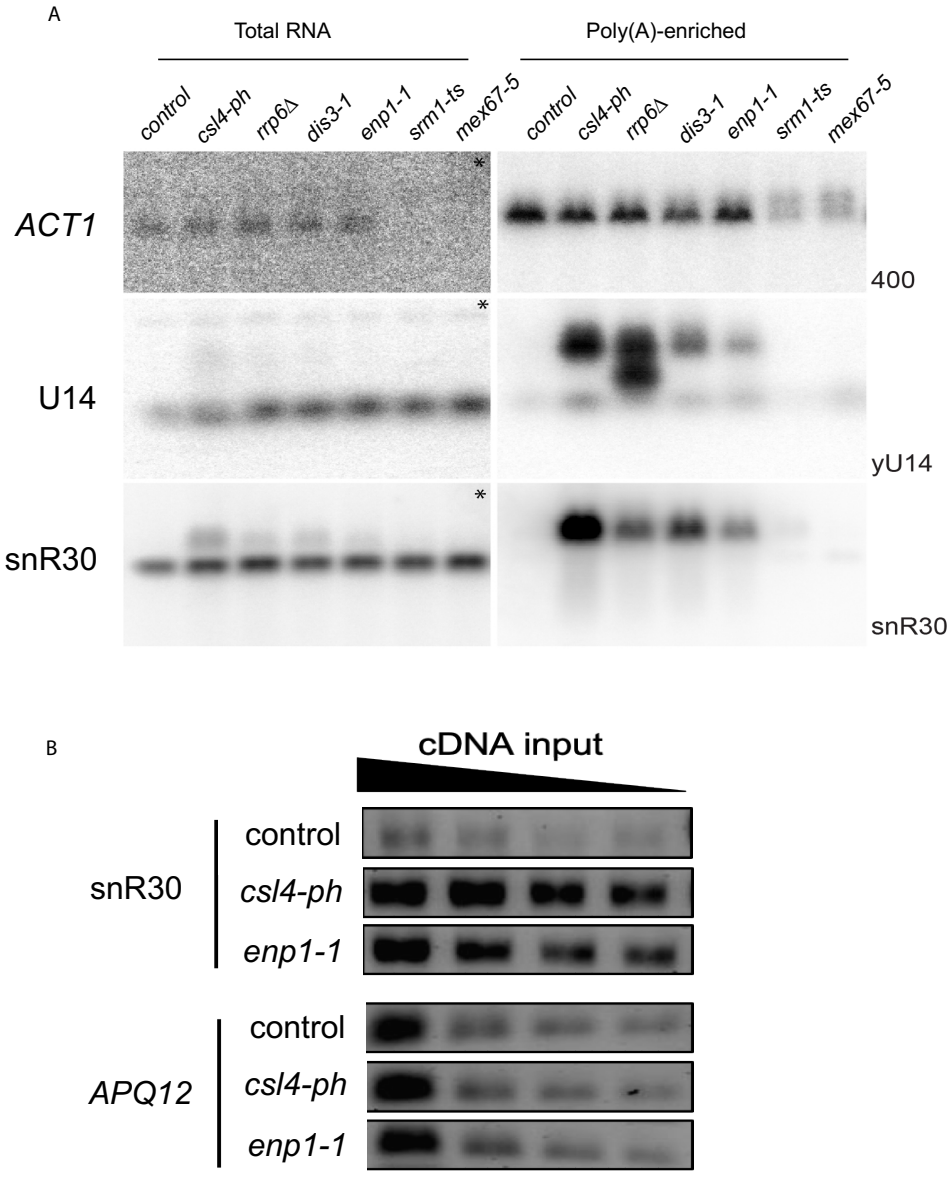


Figure 4-15. Polyadenylation of snoRNA. (A) Northern blot image of detected snoRNAs (U14 and snR30) in control, *csl4-ph*, *rrp6Δ, dis3-1*, *enp1-1*, *srm1-ts,* and *mex67-5* strains after growth at 37°C for 90 minutes from total RNA (left) and poly(A) selected RNA (right). *ACT1* mRNA served the loading control. Numbers on the right side of the panel denote probes used, sequences are listed in Table 2.4. Northern blotting was performed by Carolina Aguilar (IRCM). (B) Agarose gel image of snR30 and *APQ12* PCR products generated via an ePAT assay using total RNA isolated from control, *csl4-ph*, and *enp1-1* strains after growth at 37°C for 90 minutes. cDNA input into PCR reactions was halved for each lane from left to right.

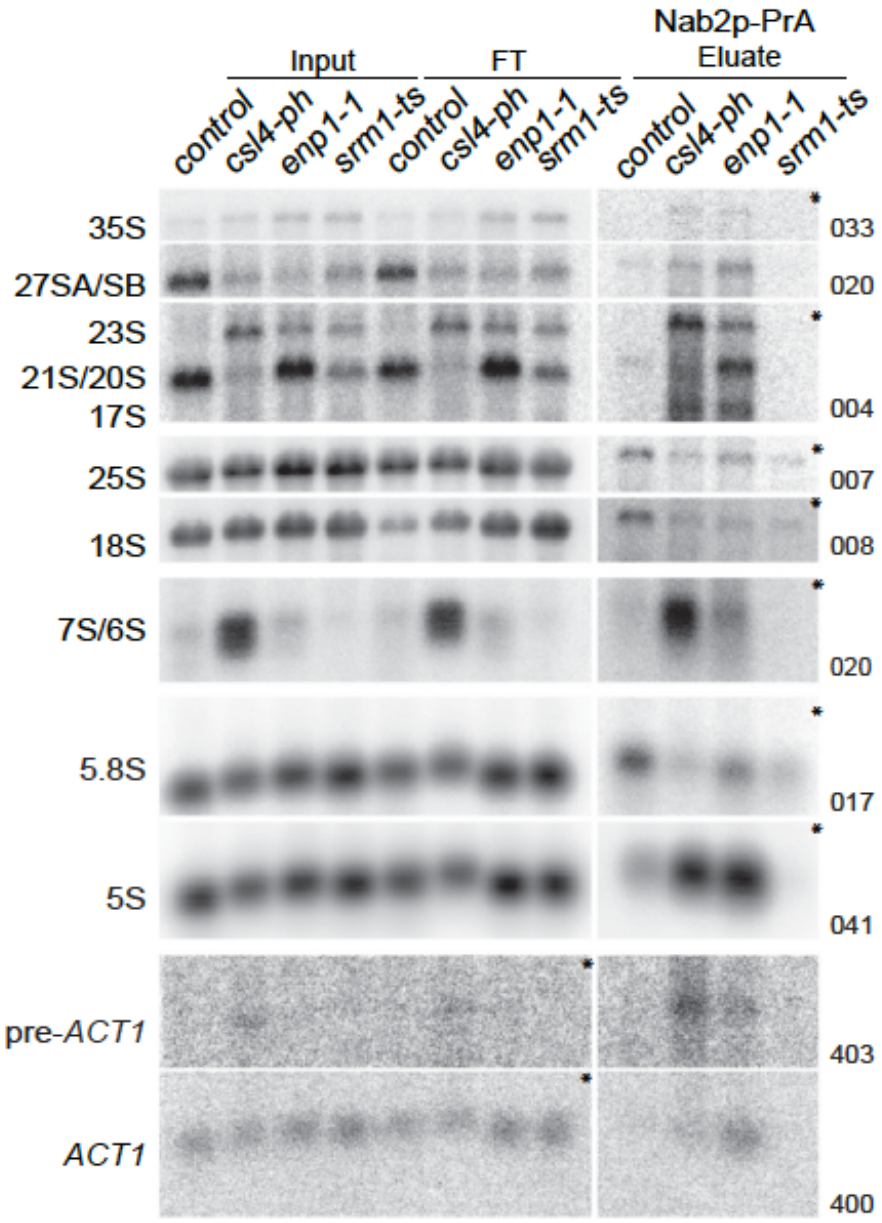


Figure 4-16. Nab2p-Protein A interactions with rRNA. Northern blot image show different pre-rRNAs (35S, 27SA/SB, 23S, 21S/20S, 17S, 7S/6S), *ACT1* mRNA, and mature rRNAs (18S, 25S, 5.8S, 5S) from control, *csl4-ph*, *enp1-1*, and *srm1-ts* mutant strains after growth at 37°C for 90 minutes. RNAs associated with Nab2p-Protein A (eluate), present in the input RNA (left), and in the flow through (FT, middle) are shown. Probes used are indicated on right of each panel, sequences are listed in Table 2.4. Northern blotting was performed by Carolina Aguilar (IRCM).

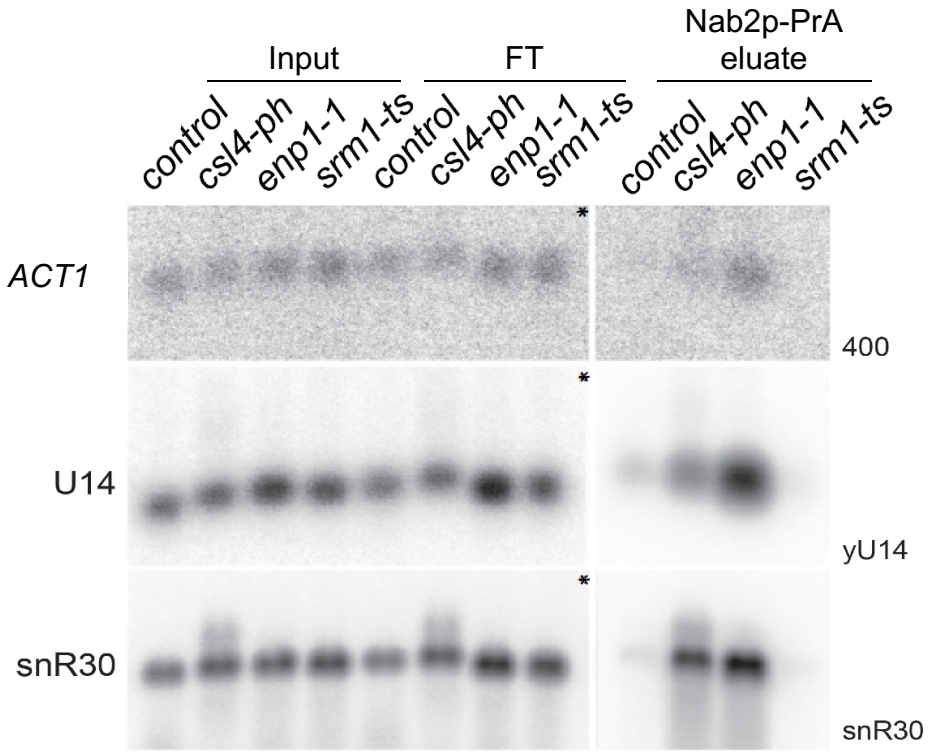


Figure 4-17. Nab2p-Protein A interactions with snoRNAs. Northern blot image showing *ACT1* mRNA, and snoRNAs (U14 and snR30) detected in control, *csl4-ph*, *enp1-1*, and *srm1-ts* mutant strains after growth at 37°C for 90 minutes. RNAs associated with Nab2p-Protein A (eluate), present in the input RNA (left), and in the flow through (FT, middle) are shown. Probes used are indicated on right of each panel, sequences are listed in Table 2.4. Northern blotting was performed by Carolina Aguilar (IRCM).

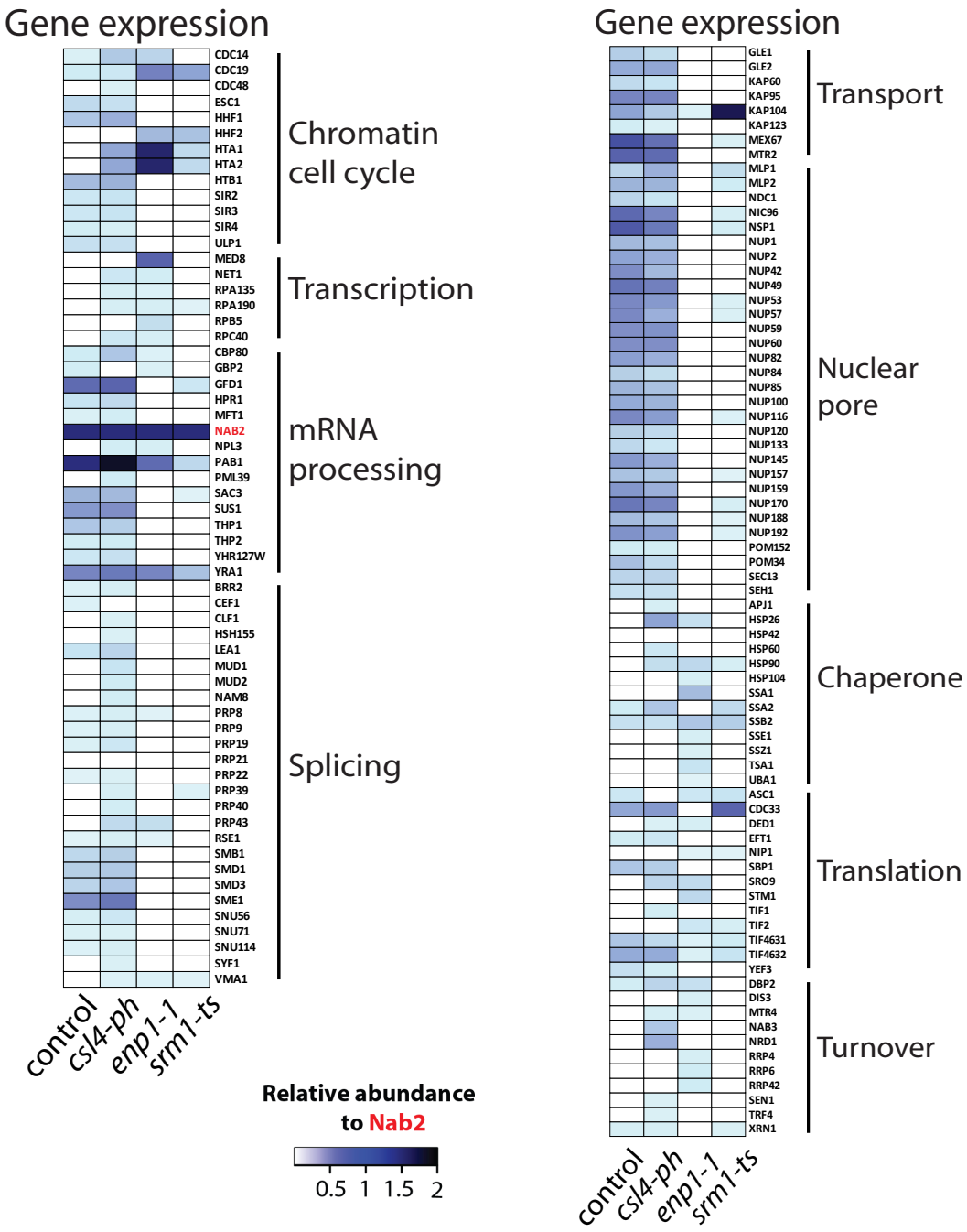


Figure 4-18. Nab2p-Protein A interactions from IP-MS analyses in *csl4-ph, enp1-1, and srm1-ts* strains. Heatmap showing relative abundance of proteins identified by AP-MS associated with Nab2p-Protein A in control, *csl4-ph, enp1-1, and srm1-ts* strains grown at 37°C for 90 minutes. Relative abundance is calculated based on the average peptide count for each protein normalized to the bait protein from three biological replicates. Proteins involved in gene expression are further categorized based on functional terms listed on the right. IP-MS carried out by Carolina Aguilar (IRCM).

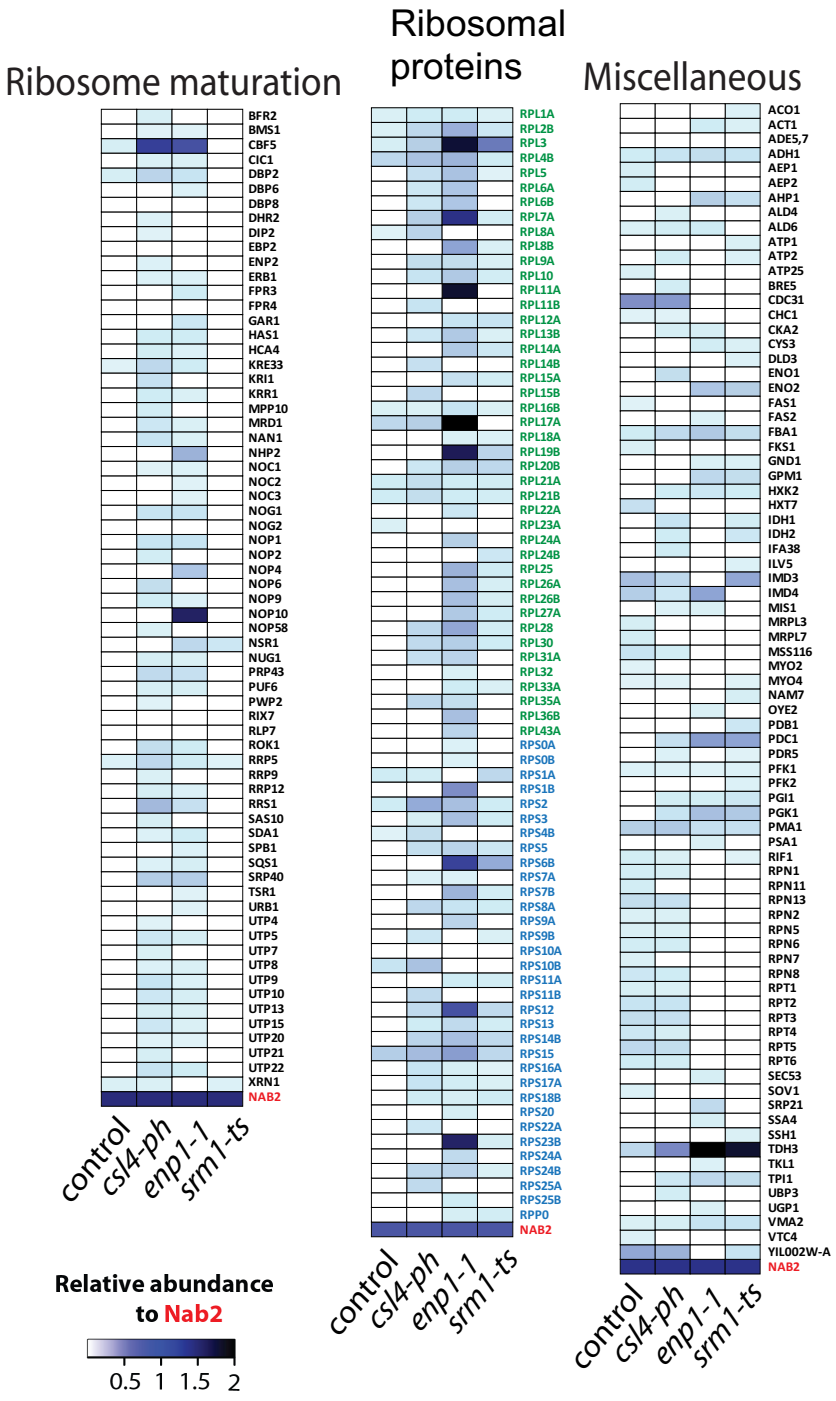


Figure 4-19. Nab2p-Protein A interactions from IP-MS analyses in *csl4-ph, enp1-1, and srm1-ts* strains. Heatmap showing relative abundance of proteins identified by AP-MS associated with Nab2p-Protein A in control, *csl4-ph, enp1-1, and srm1-ts* strains grown at 37°C for 90 minutes. Relative abundance is calculated based on the average peptide count for each protein normalized to the bait protein from three biological replicates. Proteins are categorized into ribosome maturation, green text denotes components of the large ribosomal subunit and blue text denotes components of the small ribosomal subunit, or miscellaneous. IP-MS carried out by Carolina Aguilar (IRCM).

and mRNA export pathways were almost completely absent, while increased interactions with ribosome biogenesis components were observed. This included identification of the majority of the proteins of the SSU processome and components of H/ACA-type snoRNPs (e.g. Cbf5p, Gar1p, Nhp2p, and Nop10p) that are involved in 18S rRNA processing, as well as late 40S (e.g. Krr1p, Tsr1p) and pre-60S factors (e.g. Rrp5p, Erb1p, Has1p, Noc2p, Noc3p, Rrs1p). In addition, the TRAMP associated non-canonical poly(A) polymerase Trf5p was found associated with Nab2p in *enp1-1* cells, but NNS components were not. This differed from *csl4-ph*, where Nab2p interacted with the other TRAMP associated non-canonical poly(A) polymerase Trf4p, as well as the NNS complex. This is in accordance with the reported role of Trf5p, as part of the TRAMP5 complex, preferentially performing 18S rRNA surveillance (Wery *et al.*, 2009; Choque *et al.*, 2018). Together, these data show that the Nab2p interactomes are skewed towards ribosome biogenesis machinery in both *csl4-ph* and *enp1-1* strains, in addition to related protein complexes involved in RNA processing and surveillance like the TRAMP and NNS complexes. However, differential interaction with Trf4p and NNS complex also suggest there are underlying molecular differences that drive Nab2p to engage different factors, which may reflect the initial defects caused by loss of Enp1p in ribosome biogenesis vs. the exosome more broadly in RNA processing and decay.

4.2.6 RNAPI and TRAMP activities are required for the accumulation of poly(A)-RNA and Nab2p in the nucleolus

The recruitment of RBPs to the nucleolus with poly(A)-RNA and the engagement of Nab2p with ribosome biogenesis factors, pre-rRNAs, and snoRNAs, indicate that polyadenylated ncRNA species generated in *csl4-ph* and *enp1-1* strains may be driving these phenotypes. To address if ongoing transcription is required to generate the poly(A)-RNA and RBP re-localization phenotypes, control, *csl4-ph*, and *enp1-1* strains were treated with rapamycin at the time of temperature shift. Rapamycin represses ribosome biogenesis by inhibiting TOR, which is required to stimulate rRNA biogenesis thus, rapamycin reduces transcription of rRNAs by RNAPI and ribosomal protein mRNAs by RNAP II (Loewith and Hall, 2011). By FISH, rapamycin treatment rescued the poly(A)-RNA accumulation defect in both *csl4-ph* and *enp1-1*

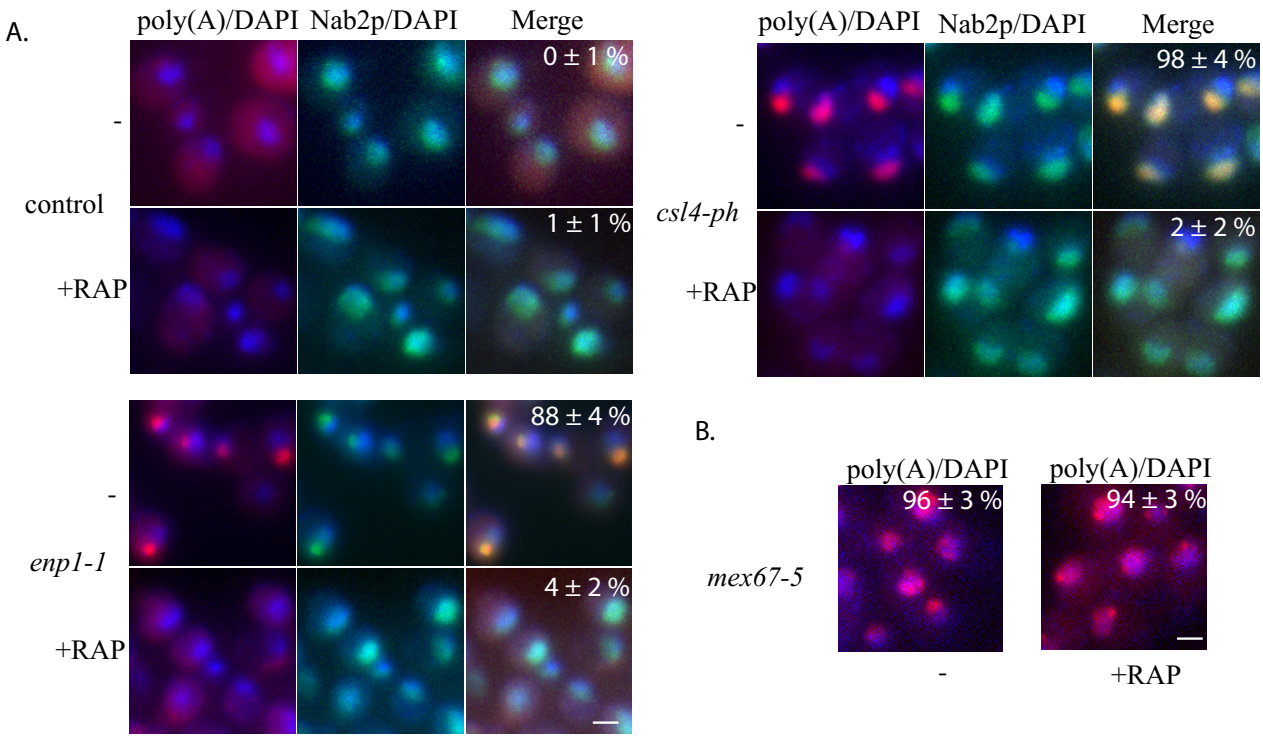


Figure 4-20. Rescue of poly(A)-RNA accumulation phenotypes during inhibition of rRNA synthesis and polyadenylation in *csl4-ph* and *enp1-1* strains. (A) Representative FISH images showing localization of poly(A)-RNA (red) and Nab2p-GFP (green) in control, *csl4-ph*, and *enp1-1* strains grown at 37°C for 90 minutes with (bottom row) and without addition of rapamycin (top row). (B) FISH images are also shown of poly(A)-RNA (red) with (right) and without (left) rapamycin addition in *mex67-5* cells grown at 37°C for 90 minutes as a control. Percentage of cells +/- SD with nuclear poly(A)-RNA accumulation are shown on the top right of the images (three replicates, > 100 cells counted for each). Scale bar = 2 μm.

strains, but it did not rescue the poly(A)-RNA defect seen in a *mex67-5* mutant, consistent with the bulk of mRNA biogenesis continuing (Figure 4-20A and B). Similarly, rapamycin treatment prevented Nab2p-GFP re-localization to the nucleolus in *csl4-ph* and *enp1-1* strains (Figure 4-20A and B). Rescue of both the poly(A)-RNA accumulation and Nab2p re-localization suggested that it is the accumulated polyadenylated pre-rRNAs that drive the development of the Nab2p relocalization phenotype observed in these mutants.

Given that pre-rRNAs appear to be central to the development of the poly(A)-RNA accumulation phenotype, it follows that this would be dependent on TRAMP activity, since TRAMP targets rRNAs to mediate aspects of their biogenesis (de la Cruz *et al.*, 1998; Tudek *et al.*, 2014; Thoms *et al.*, 2015). A prediction of this is that depletion of the TRAMP poly(A)-polymerase activity would lessen or rescue the poly(A)-RNA accumulation phenotype. Indeed, deletion of either of the genes encoding the TRAMP associated non-canonical poly(A)-polymerases, *TRF4* or *TRF5,* resulted in a reduction in the level of poly(A)-RNA detected in *enp1-1* cells after 90 minutes at 37°C (Figure 4-21). In the case of *csl4-ph*, double mutants with *trf4Δ* were not obtained, but *csl4-ph-trf5Δ* mutants similarly accumulated far less poly(A)-RNA in the nucleolus (Figure 4-21). These findings suggest that ongoing ribosomal RNA transcription and processing by the TRAMP complex is required for the nucleolar poly(A)-RNA accumulation.

4.2.7 Excess poly(A)-RNA is toxic to the cell

A model that begins to emerge from these data is that alterations in total nuclear poly(A)-RNA levels through altered rRNA biogenesis feeds back on and disrupts other RNA processing pathways, involving ncRNAs, mRNAs, and pervasive transcripts. In these cases, loss of Csl4p or Enp1p activity leads to the accumulation of polyadenylated ncRNAs that act as a competitive 'sink' for RBPs, including Nab2p, which perturbs nuclear RNA processing, and this has a significant impact on the cellular transcriptome. Poly(A)-RNA binding proteins are likely central to the development of these phenotypes given the inherent non-specific binding of PABPs to a poly(A) sequence. In this way, increasing levels of poly(A)-RNAs in the nucleus would function to titrate PABPs and associated factors away from other RNA biogenesis functions, leading to further errors in RNA biogenesis, the generation of more diverse classes of exosome substrates,

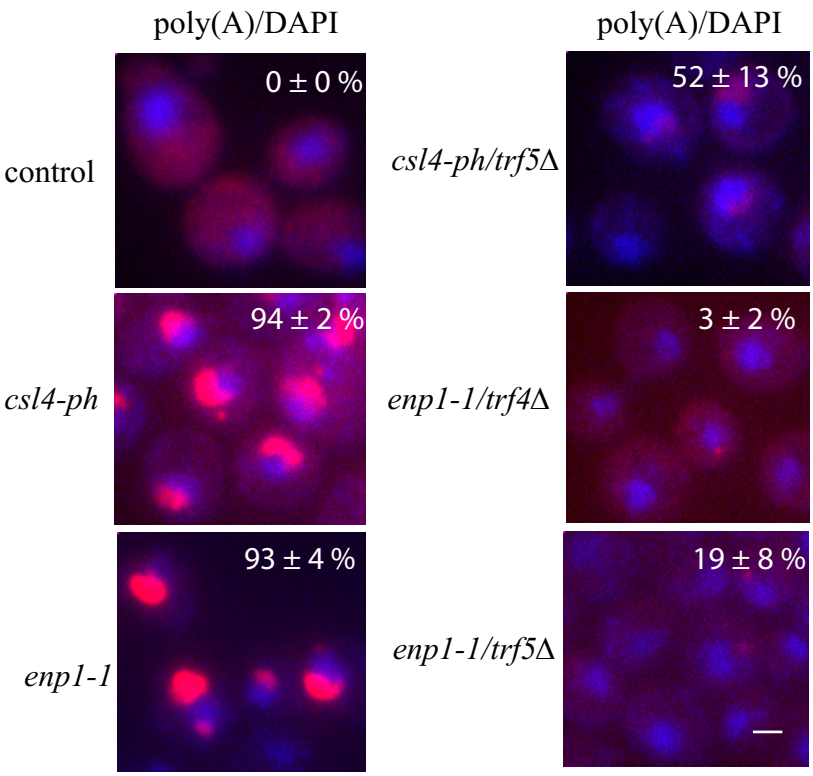


Figure 4-21. Rescue of poly(A)-RNA accumulation phenotype during inhibition of polyadenylation in *csl4-ph* and *enp1-1* strains. Representative FISH images showing localization of poly(A)-RNA (red) in comparison to the DAPI stain (blue) in control, *csl4-ph, enp1-1, csl4-ph / trf5Δ, enp1-1 / trf4Δ* , and *enp1-1 / trf5Δ* strains grown at 37°C for 90 minutes. Percentage of cells +/- SD with nuclear poly(A)-RNA accumulation are shown on the top right of the images (three replicates, > 100 cells counted for each). Scale bar = 2 μm.

the accumulation of more RNPs within the nucleus, and thus further dampening both nuclear RNA processing and nuclear surveillance activities. In this scenario, the overexpression of a poly(A)-RNA binding protein could buffer this positive-feedback loop and limit development of the observed phenotypes. To test this possibility, the poly(A)-RNA binding protein Pab1p was overexpressed in control and mutant strains to provide excess PABP activity that can bind and compete for poly(A) tails. After 90 minutes at 37°C, the overexpression of Pab1p was found to prevent the development of a poly(A)-RNA accumulation phenotype in an *enp1-1* strain, but not the phenotype of a *csl4-ph* or *mex67-5* strain (Figure 4-22A). Given that the exosome is responsible for the degradation of many classes of RNAs, including pervasive transcripts, and TRAMP's polyadenylation activity likely continues, it is not surprising that overexpression of Pab1p does not rescue a *csl4-ph* mutant. Rescue of an mRNA export mutant is similarly not expected. Importantly, the rescue of *enp1-1* is as predicted by the PABP titration model. Cells that overexpressed Yra1p, an RBP that is also recruited to the poly(A)-RNA mass in *csl4-ph* and *enp1-1* strains, but is not a PABP, retained the poly(A)-RNA accumulation phenotype in both *csl4-ph* and *enp1-1* strains (Figure 4-22A). Finally, if excess poly(A)-RNA is the catalyst of these defects, it would be expected that overexpression of a PABP may improve the overall fitness of *enp1-1* cells at a semi-permissive temperature. Indeed, overexpression of Pab1p significantly improved the growth of *enp1-1* at 32°C, while having no impact on a control or a *mex67-5* strain (Figure 4-22 B & C). Overall, these data provide strong evidence that the accumulation of polyadenylated material in *enp1-1* strain generated through a disruption to rRNA biogenesis is toxic to the cell through, at least in part, by competing for PABPs.

4.3 Discussion

Noncoding and messenger RNA processing is mediated by RBPs that function to support nuclear RNA biogenesis. The findings presented here provide evidence that excess nuclear poly(A)-RNA is toxic to the cell, disrupting RNA processing through sequestering RBPs that function in mRNA biogenesis and export, including Nab2p. Moreover, convergence on a shared set of phenotypes in mutants that function in relatively distinct biological processes provides

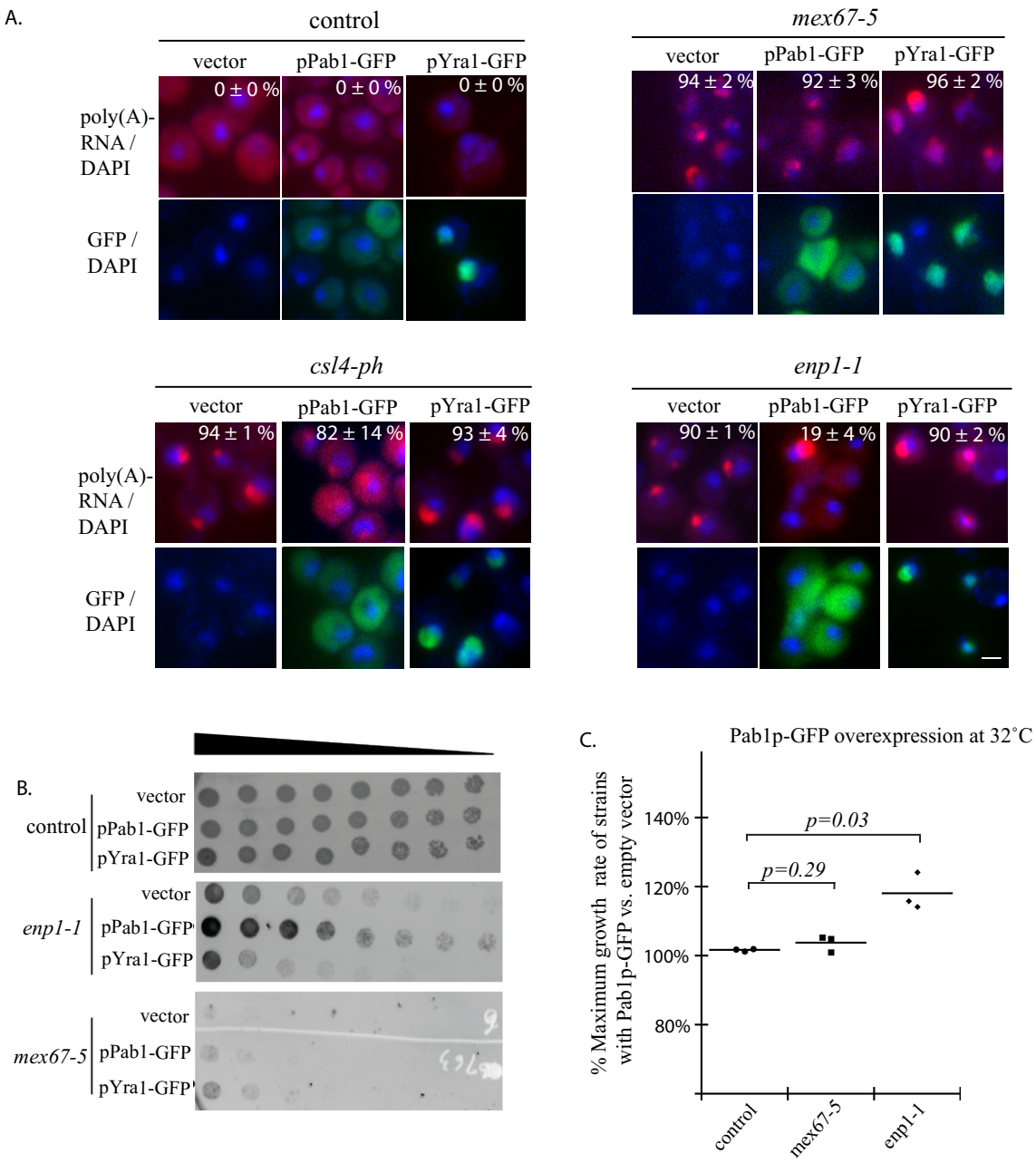


Figure 4-22. Rescue of the poly(A)-RNA accumulation and growth defects in an *enp1-1* strain through overexpression of the PABP Pab1p. (A) Representative FISH images showing localization of poly(A)-RNA (red) in control, *csl4-ph*, *enp1-1*, and *mex67-5* strains overexpressing Pab1p-GFP (green) after growth at 37°C for 90 minutes as compared to the empty plasmid control. Percent of cell with poly(A) are given on the FISH images (three replicates, > 100 cells counted for each). Scale bar = 2 μm. (B) Growth of control, *enp1-1*, and *mex67-5* strains at 32°C when overexpressing Pab1p-GFP. From the first spot, each subsequent spot is a two-fold dilution from left to right. Image was taken after three days of growth. (C) Growth rates of control, *enp1-1*, and *mex67-5* strains at 32°C overexpressing Pab1p-GFP normalized to the growth rate of the same strains carrying an empty plasmid. P-value calculated using a two-tailed Students t-test.

evidence that multiple initiating events with unique molecular origins can ultimately lead to a set of shared terminal phenotypes. This suggests that numerous cellular changes, including excess poly(A)-RNA production, the inability to export poly(A)-RNA from the nucleus, decreased levels of nuclear PABPs or RNA processing factors, and the inability to decay poly(A)-RNA could all initiate events that lead to an overall loss of nuclear RNA homeostasis.

Polyadenylation has long been associated with mRNA processing leading to export, translation, and transcript stability. These cellular processes are mediated in part by the PABPs that are associate with a the poly(A) tail (Tudek *et al.*, 2018a). With time, it has become evident that polyadenylation also occurs on ncRNAs, including rRNAs, snoRNAs, and pervasive transcripts, often directing RNA processing or decay by the exosome (Kuai *et al.*, 2004; Kim *et al.*, 2006). This raises the issue of specificity, i.e. how is it that a poly(A) tail on an mRNA and ncRNA are able to be distinguished to direct differing outcomes? Context is clearly important, with events upstream of polyadenylation leading to specific RNP architectures (e.g. protein constituents and structures), that when combined with features like the length of the transcript, the polymerase being used, and associated polyadenylation machinery, would lead to specific outcomes. For example, mRNAs associated with the NNS complex, as compared to the CPF complex, would be more likely to be delivered to the exosome for decay (Kim *et al.*, 2006; Tudek *et al.*, 2014). Whereas, if an mRNA is bound by the CPF, the transcript is likely to be processed and exported to the cytoplasm (Brockmann *et al.*, 2012). However, it is expected that the fate of polyadenylated RNAs are influenced by all nuclear machineries that can engage the transcript. The exact outcome is the result of the affinities between all the factors involved, the sub-nuclear localization and availability of each component, and kinetics of the competing processes. Moreover, it would be expected that the outcome would change with cellular state, with even small changes in mRNA or ncRNA processing efficiencies influencing one another. Speculatively, this may allow cells a mechanism to coordinate changes in mRNA and ncRNA transcript processing during periods of rapid changes in gene expression (e.g. stress).

Here it is reported that in the case of a *csl4-ph* mutant, the core exosome is largely intact, but loses contact with the catalytic subunit Rrp6p and the TRAMP complex, thus impeding its function (Zinder and Lima, 2017). Despite this impairment of the exosome, nuclear/nucleolar exosome substrates are still tagged with short poly(A) tails by the TRAMP complex (Vanácová *et al.*, 2005), thus creating a rapidly increasing pool of PABP targets that can sequester RBPs such

as Nab2p in the nucleolus. Notably, these perturbations to mRNP biogenesis through the sequestration of RBPs and generation of surveillance substrates could further lead to other RNA processing complexes being redistributed, including the NNS complex (Moreau *et al.*, 2019). As a result, other NNS-dependent processes, including the surveillance of pervasive transcripts and snoRNA 3′ processing may be altered. Notably, these are defects observed in an *enp1-1* mutant. This is again suggestive of a homeostatic relationship between all nuclear RNA processing pathways that could lead to coordination of RNA processing, but also collapse due to competition over of a limited pool of nuclear components (Figure 4-23).

In support of this, here it was demonstrated that altering nuclear import, rRNA biogenesis, or exosome activity caused the formation of excess poly(A)-RNA that is toxic to the cell. The results of excess nuclear poly(A)-RNA is the sequestration of PABPs and other RBPs, which in turn results in further failures in RNA processing, forming a positive feedback loop that drives cells towards a set of shared terminal phenotypes, including a common cellular transcriptome defined by stabilized pervasive transcripts and polyadenylated ncRNAs. These phenotypes reflect the impact of excess poly(A)-RNA and the associated disruption in nuclear homeostasis on the cell, not the function of the disrupted gene product. This argument is supported by the finding that these phenotypes could be suppressed by lowering ncRNA production through disruption of RNAPI transcription, disrupting TRAMP polyadenylation activity, or by overexpressing a PABP that could buffer this effect by competing for poly(A)-RNA. These modes of rescue all highlight the need for the cell to maintain poly(A)-RNA levels in balance with PABPs.

Other recent work has shown similar relationships, including the finding that depletion of Nab2p by induced degradation leaves nascent transcripts vulnerable to nuclear decay (Schmid *et al.*, 2015). Similarly, another publication showed that when mRNA export is blocked, accumulated mRNAs sequester Nab2p leaving nascent mRNA unprotected and subject to decay (Tudek *et al.*, 2018b). In this thesis, down regulation of most cellular mRNAs was observed in the *dis3-1* and *csl4-ph* exosome mutants, which appears to counter the Nab2p versus exosome model, since loss of exosome activity would be expected to promote mRNA stability. However, based on data presented here, the requirement of the exosome for the processing of highly abundant ncRNAs and pervasive transcripts leads, over time, to high levels of stabilized nuclear ncRNAs that sequester PABPs and other shared RBPs. The depletion of RBPs ultimately causing decreased RNAPII

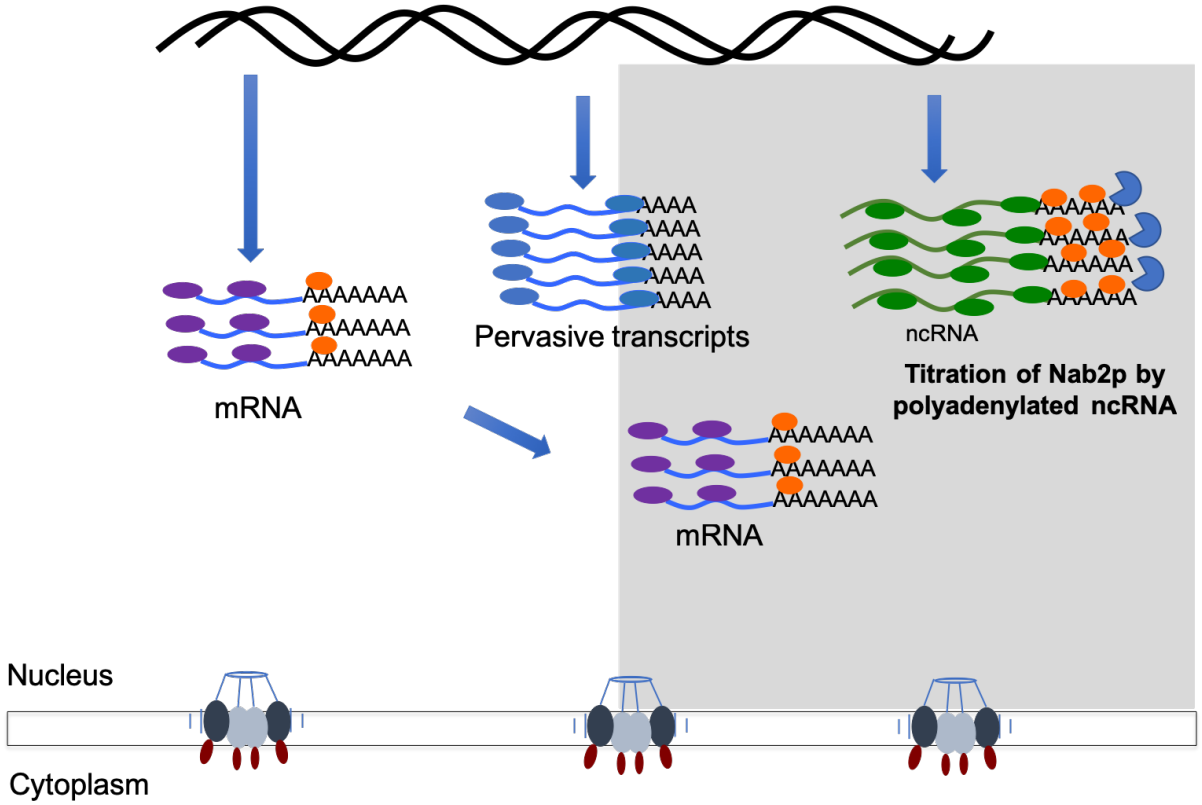


Figure 4-23. Model for disruption of nuclear RNA homeostasis in *enp1-1*. Disruption of ncRNA processing leads to accumulation of polyadenylated ncRNAs in the nucleolus (gray). Polyadenylated ncRNA binds mRNA binding proteins, specifically PABPS, which titrate these factors away from normal mRNA processing pathways. Sequestration of mRNA processing factors (e.g. Nab2p) leads to further errors in mRNA biogenesis, redistribution of RBPs, a genome wide down regulation of mRNA levels, and a disruption to mRNA export. In addition, excess polyadenylated ncRNAs could saturate the nuclear exosome and disrupt exosome function leading to the stabilization of pervasive transcripts.

transcription rates and/or efficiencies and a reduction in mRNA transcripts.

In addition, while core exosome activity may be largely absent due to loss of Csl4p or Dis3p, it is possible that nascent transcripts are degraded by Rrp6p via exosome independent mechanisms, which have been reported (Callahan and Butler, 2008). Retention of transcripts in the nucleus and their lack of protection due to inadequate binding of RBPs may further promote decay by providing time for Rrp6p to find and degrade these unprotected transcripts. In line with this close association between Nab2p and Rrp6p, it has been shown that deletion of *RRP6* stabilizes unprotected nuclear RNA and rescues growth defects in a *nab2-1* strain (González-Aguilera *et al.*, 2011). Even more surprisingly, deletion of *RRP6* has been reported to render Nab2p non-essential, as a *rrp6Δ-nab2Δ* double deletion mutant was found to be viable (González-Aguilera *et al.*, 2011; Schmid *et al.*, 2015). As such, the number of processes supported by the exosome, with the overall complexities of nuclear processing requiring shared machineries and RBPs, provides for a scenario where transient loss of functions may favour one process (e.g. mRNA stabilization) over another, but long-term this likely leads to an overall reduction in RNA biogenesis and common set of failures in nuclear RNA processing.

The concept of nuclear RNA processing homeostasis and the idea that a failures in RNA processing in one pathway can feedback on others is best exemplified by the *enp1-1* mutation, since unlike the exosome, Enp1p has only been linked to the processing of the 18S rRNA. Upon loss of Enp1p, rRNA processing stalls, causing stabilization of longer 18S pre-cursors that can be polyadenylated (23S, 21S, and 20S, see Figure 4-14). Why these precursors are disruptive to global RNA processing is not clear, but at least two major possibilities exist. First, these precursors are produced in large amounts, have high affinity for the exosome complex, and/or are difficult to process, leading to competitive inhibition of the exosome, which ultimately phenocopies an exosome mutant (Figure 4-23). A second possibility is that these pre-cursors are not recognized or processed by the exosome due to overall changes in RNP structure, leading to the build-up of these precursors and sequestration of RBPs normally involved in rRNA processing, or in other RNA processing events, like Nab2p. This may limit the activity of these RBPs, leading to errors in other RNA processing pathways, more substrates to be cleared by the exosome and eventual saturation of exosome surveillance capacity, leading to more failures, and a terminal phenotype that is similar to an exosome mutant. Whether either of these models accurately describes what is occurring in an *enp1-1* strain requires further investigation.

Chapter V: *Perspectives*

5.1 Synopsis

The lines of questioning and resulting data presented in Chapters III and IV were built upon a functional screen for essential genes involved in mRNA export. This work, together with the results of another screen carried out by the Silver Lab (Harvard Medical School) focused on non-essential genes, provides important information regarding poly(A)-RNA accumulation phenotypes that result from the perturbation of the majority of yeast genes (Casolari *et al.*, 2004; Paul and Montpetit, 2016). Still, the budding yeast *Ts* collection does not contain every essential gene in yeast, thus essential mRNA export factors may still remain to be discovered (Ben-Aroya *et al.*, 2008; Li *et al.*, 2011). Moreover, these screens were all performed in rich media under logarithmic growth, so it is expected that during stress or other growth stages (e.g. stationary phase or meiosis) other factors may become essential for mRNA export. These facts suggest that there are still factors to be discovered that function in these specialized cases. However, it is expected that we now have a near complete list of the core factors that function in mRNA export, which can now be used to build and inform current models of mRNA export.

Most previous works have characterized perturbations in mRNA export mutants based on bulk poly(A)-RNA accumulation phenotypes. Notably, these studies lack information regarding that nature of the poly(A)-RNA material (i.e. mRNA vs. ncRNA) and do not address if the mutated factor functions broadly in mRNA export or on a small subset of mRNAs. Importantly, hits from the screen presented in Chapter III were characterized as mRNA export mutants based on assays that extended beyond poly(A)-RNA accumulation. Specifically, further work was carried out to determine the localization patterns of individual mRNAs, including transcripts that are spliced, un-spliced, and have different RBP binding profiles. Through these analyses, it was determined that not all mutants that accumulate poly(A)-RNA in the nucleus impact mRNA export in a similar manner. For example, mutations in gene products that are known to be directly involved in mRNA export accumulated mRNAs at the nuclear envelope, while mutations in gene products that affect mRNA 3′ end processing and elongation accumulated mRNAs at transcription sites. Yet, both of these classes of mutants showed a very similar accumulation of bulk poly(A)-RNAs in the nucleus. A third class of mutants altering ncRNA processing and nuclear surveillance showed both poly(A)-RNA and mRNA accumulation in the nucleolus; however, from follow up work (Chapter IV), it is clear that some amount of the poly(A)-RNA

signal in these mutants originates from stably polyadenylated ncRNAs. These findings demonstrate the importance of single molecule mRNA FISH to identify and better characterize mRNA export mutants, and to determine the step of nuclear mRNA processing that a factor may function within. Future follow up work could include re-evaluation of known mRNA export mutants using single molecule FISH to characterize the stage of nuclear processing being impacted by a given mutant. It is expected that such work would provide data to order factors within the nuclear mRNA processing pathways and to detail whether the factor functions on all or a subset of mRNAs.

As discussed above, the aim of the screen presented in Chapter III was to identify essential genes involved in mRNA export, but the majority of the novel mutants identified in this study appear to indirectly affect mRNA export rather than directly function as mRNA export factors. This again suggests that most factors in the export pathway are known, but also indicates connections between mRNA export and other nuclear processes. For example, the largest group of hits included genes that are involved in chromosome segregation. Future research will be required to explore these links between poly(A)-RNA processing, RNA export, and chromosome segregation (see section 5.2). As mentioned above, another class of mutants were identified in which poly(A)-RNA accumulated specifically in the nucleolus. This class included mutants that impacted ncRNA processing and nuclear exosome function. The work presented in Chapter IV specifically focused on this set of mutants and the functional link between ncRNA processing and nuclear RNA homeostasis, which appears to be maintained in part through a balance in the functions of the nuclear exosome complex and the poly(A)-binding protein Nab2p. In section 5.3, these findings are discussed in the context of other literature linking nuclear RNA homeostasis to PABPs and exosome function. Future directions for investigations are also presented that could provide insight into important mechanistic details surrounding the regulation of nuclear poly(A)-RNA levels and its impact on overall gene expression and RNA processing.

5.2 The links between chromosome metabolism and mRNA processing

The eukaryotic cell possesses distinct sets of proteins involved in mediating the process of chromosome cohesion, condensation, and segregation during cell division. This includes

kinetochore proteins, components of the mitotic spindle and spindle pole body, and DNA bound factors that help shape and organize chromosomes (Palmer *et al.*, 1991; Schneider and Grosschedl, 2007). In the screen presented here aimed at identifying novel mRNA export factors (Chapter III), genes functioning in each of these categories were identified as causing a poly(A)-RNA accumulation defect when mutated. This included Ipl1p, Spc24p, Sli15p, Cep3p, and Mps1p, which are each involved in chromosome segregation, as well as Smc1p, Smc3p, and Smc4p, proteins involved in chromatid cohesion and condensation (Paul and Montpetit, 2016). It is unlikely that these factors have a direct role in mRNA export, but some of these factors have also been linked to disease (Krantz *et al.*, 2004), making these phenotypes of interest given their impact on RNA processing when mutated. For the factors engaged in chromosome segregation, one explanation is that mutations causing errors in chromosome segregation result in the random loss of a chromosome, or part of a chromosome, containing an essential mRNA export factor and this ultimately results in an mRNA export defect. This idea is supported by the fact that only a small percent (20%) of the population of cells developed a poly(A)-RNA accumulation phenotype in these mutants. In line with this, our work showed that in chromosome segregation mutants, signals from rRNA processing intermediates (*ITS*) and a tagged nucleolar protein Nop56p-GFP were lost in cells with a poly(A)-RNA accumulation defect. This may be due to loss of the rDNA region, which is segregated late in budding yeast and may be prone to high rates of loss (Machín *et al.*, 2005). As such, the factor being lost could be a gene situated near the rDNA locus. Alternatively, recent research suggests that upon mis-segregation of the nucleolus the majority of exosome complex is also mis-segregated to one of the daughter cells due to the enrichment of the exosome in the nucleolus (Hocquet *et al.*, 2018). If this is the case, the mechanism for poly(A)-RNA accumulation in these mutants may result from a breakdown in nuclear homeostasis that is caused by altered exosome activities. Future experiments aimed at describing which DNA regions are lost in these mutants would be useful in deciphering these details, as would further details surrounding exosome activity in the resulting aneuploid cells.

In the case of the Smc proteins, these factors form a ring like protein complex that functions to link two DNA molecules, which is central to organizing and segregating the genetic material (Nasmyth and Haering, 2005; Hirano, 2006; Haering *et al.*, 2008). By regulating the structure of chromosomes within the nucleus, the Smc proteins also play an important role in gene expression (Dowen and Young, 2014). For example, Smc proteins function in silencing of

the mating type genes in yeast, and by a similar mechanism, Smc proteins also impact the silencing of tRNA genes and artificially introduced genes at heterochromatic pericentric DNA in *S. pombe* (Bhalla *et al.*, 2002; Iwasaki *et al.*, 2010; He *et al.*, 2016; Wang *et al.*, 2016). In chicken DT40 cells, condensin I was found to be associated with active promoters in M phase, with depletion of condensin I leading to reduced expression of a subset of genes in G1 (Kim *et al.*, 2013). Cohesin and condensin II are further known to reduce expression of genes driven by super-enhancers in embryonic stem cells, with depletion of either impacting gene expression and cell identity (Dowen *et al.*, 2013). Given these Smc protein functions, it is possible that the accumulation of poly(A)-RNA occurs for the reasons described above (i.e. segregation errors); however, it is also possible that loss of Smc activity induces structural changes within chromosomes that cause alterations in gene expression (Dowen *et al.*, 2013; Dowen and Young, 2014). For example, it may be that upon loss of Smc activity, the transcription, processing, and export of some transcripts are impacted due to perturbations in chromosome organization. What genes are impacted and how this impacts RNA processing and mRNA export requires further investigation of these possible relationships. These impacts could come from alterations in the 3D structure of chromatin itself in SMC mutants, altered distributions of RNAPII, or changes in gene expression in a gene important for mRNA export. Future studies aimed at determining gene expression or RNAPII distributions across the genome of these mutants may allow such questions to be addressed. Finally, Smc proteins are known to be organized around transcriptional units (Glynn *et al.*, 2004), making it possible that the function of these proteins is much more direct; as such, the study of the interactions between Smc factors, RNA, and RNA binding proteins may be an enticing area for future study.

5.3 The role of the nucleolus on the mRNA life cycle

The nucleolus is the center for rRNA and snoRNA biogenesis; however, its role in mRNA biology is not well studied. Interestingly, work carried out in Chapter III has found that mRNAs accumulated in the nucleolus when exosome activity was impaired, or when rRNA processing was altered. This was not gene specific, as every mRNA tested was found to accumulate in the nucleolus of *epn1-1* and exosome mutants, which may suggest a general role for the nucleolus in

nuclear mRNA processing. What that role may be is not known, but many interesting possibilities exist. For example, an mRNA might transit through the nucleolus for processing or modification, which may be essential for export or specific events in the cytoplasm. Indeed, evidence does exist for mRNA pseudouridylation (Carlile *et al.*, 2014), which would likely occur in the nucleolus, and for specific mRNAs transiting through the nucleolus to facilitate cytoplasmic localization (Du *et al.*, 2008). Still, a well characterized or general role of the nucleolus in the formation of export ready mRNAs has not been described. Therefore, further studies using live-cell mRNA imaging approaches may be an important starting point, which have previously been used to address mRNA export (Smith *et al.*, 2015). Applying such technology in the context of a nucleolar marker would allow one to define the frequency of an mRNA entering the nucleolus, the amount of time spent there, and where the mRNA travels after the nucleolus. In the context of a ncRNA processing mutant, the observed accumulation of mRNAs in the nucleolus could be the result of a block in nucleolar mRNA processing and export due to changes in nucleolar structure or function. However, an alternate possibility is that upon loss of nucleolar homeostasis, ongoing errors in mRNA processing lead to defective mRNAs accumulating that are destined for decay as part of a quality control mechanism. In the latter case, mRNAs that transit to the nucleolus would never be exported to the cytoplasm, but targeted to the nucleolus to be decayed by the nuclear exosome. Single molecule tracking experiments may also provide such information through the observation of mRNAs that go to the nucleolus but are never exported to the cytoplasm. It is expected that distinguishing such possibilities will be important, as it would define this compartment as either being associated with proper mRNA maturation and export, or quality control and mRNA decay.

5.4 The buffering of mRNA levels in the nucleus

In chapter IV evidence is presented that shows mRNA levels are reduced in mutants that accumulate poly(A)-RNA in the nucleolus. In response to poly(A)-RNA accumulation, a recent report by the Jensen Lab (Aarhus University) suggests that Nab2p activity becomes limiting and transcriptome levels change due to Rrp6p mediated decay of mRNA with unprotected poly(A) tails (Schmid *et al.*, 2012). This raises the possibility that the lower levels of mRNA observed in

the mutants tested here are the result of Rrp6p mediated decay of nascent mRNAs due to the sequestration of Nab2p, and other RBPs, by polyadenylated ncRNAs. In line with this close relationship between Rrp6p and Nab2p, it has been reported that a *nab2Δ* strain is viable in the presence of *rrp6Δ* (González-Aguilera *et al.*, 2011). As such, further studies could focus on Rrp6p independent functions in exosome core mutants to address if Rrp6 is still active and can associate with nascent mRNAs for their degradation. In fact, it is possible that in mutants that have Rrp6p disengaged from the exosome core (e.g. *csl4-ph*) that Rrp6p could be more active against exosome-independent targets, including nascent mRNAs. An idea that could be tested by performing cross-linking immunoprecipitation (CLIP) experiments in *csl4-ph* mutant to map Rrp6p binding to transcripts. Recent research has also shown that cytoplasmic decay factors, e.g. Xrn1p, travel to the nucleus in order to regulate transcription in a nuclear RNA decay mutant (Sun *et al.*, 2013). Thus, it is possible that Xrn1p activity could also be a cause of lowered mRNA levels in the mutants we have tested. A ChIP-seq experiment with Xrn1p in *csl4-ph* or *enp1-1* could be performed to test whether Xrn1p localization increases at the loci of downregulated genes. A final explanation for the downregulation of mRNAs is that factors needed for transcription are less abundant due to sequestration on nucleolar mRNAs and ncRNAs, which leads to decreased RNAPII activity. A hypothesis that could be addressed by NET-Seq or RNAPII ChIP-seq to address ongoing transcription and RNAPII occupancy at gene loci.

5.5 The role of Enp1p in rRNA processing and nuclear RNA homeostasis

The collection of *Ts* strains screened in Chapter III contain > 20 rRNA processing mutants, yet the screen only identified *enp1-1* as accumulating poly(A)-RNA and mRNA in the nucleolus. In Chapter IV, it was further shown that the *enp1-1* strain accumulated both 20S and 23S polyadenylated pre-rRNAs. This raises questions relating to the differences in phenotypes arising from losing Enp1p function as compared to other rRNA processing factors, since it is well known that other rRNA processing mutants also accumulate pre-rRNAs in the nucleus (Chen *et al.*, 2003). In other words, what makes Enp1p different? Potentially, other processing mutants may arrest rRNP maturation at a stage where the rRNP composition or structure is easily decayed by the exosome. In contrast, in *enp1-1*, the rRNP structure may not be efficiently recognized by

the exosome complex or may be too stable for the exosome to effectively decay the RNA within the rRNP. CLIP-Seq approaches may be one approach that could be taken to discern if the exosome is engaging stalled rRNPs in an *enp1-1* mutant and where this is occurring. Another possible hypothesis is that Enp1p itself is required for the exosome to target the polyadenylated rRNA. However, Rrp41p-Protein A interaction data from a control strain did not show any evidence of an interaction between Enp1p and the core exosome, but without crosslinking or otherwise stabilizing protein-protein interactions, the overall level and transient nature of the interaction may make this hard to detect. Ultimately, the question of what happens in an *enp1-1* mutant would probably be best addressed through a structural approach, which should be methodologically possible given the many recent crystal and cryo-EM structures that have been generated for rRNPs in yeast (Greber *et al.*, 2012; Schmidt *et al.*, 2016; Brown and Shao, 2018; Sosnowski *et al.*, 2018). A final hypothesis is that the late nuclear role for Enp1p in promoting the export of mature 20S rRNAs results in the accumulation of rRNPs that are not recognized by the exosome as being defective, as such they readily accumulate in the nucleus. In these cases, nuclear accumulated 20S rRNA would undergo TRAMP mediated polyadenylation, and being the most highly abundant RNA produced in the cell, these rRNAs would begin to compete for both the nuclear surveillance machinery, as well as any RBPs they can engage, disrupting nuclear RNA processing homeostasis.

5.6 Closing Remarks

Overall, this work provides new insights into the interconnections between stable ncRNA processing and mRNA expression. The exosome and certain RBPs acting as key components that functionally connect ncRNA processing and mRNA expression through competition for these factors under normal or perturbed cellular conditions. More broadly, it is expected that these concepts can aid our understanding of how human disease results through mutations in core RNA processing components, like the exosome. For example, mutations in the RNA exosome genes EXOSC3, EXOSC8, and EXOSC2 (*RRP40*, *RRP43* and *RRP4* in *S. cerevisiae*) cause tissue-specific disease phenotypes (Morton *et al.*, 2018). Mutations in exosome cofactors also cause tissue-specific diseases of varying phenotypes (Hartley *et al.*, 2010; Fabre *et al.*, 2012; Giunta *et*

al., 2016). Similarly, mutations in the PABP ZC3H14, a Nab2p ortholog, are associated with inherited intellectual disability (Pak *et al.*, 2011). Currently, how mutations in these RNA surveillance and processing factors result in these particular diseases is not known. Here it is demonstrated that generation of excess or improperly processed RNA substrates can alter RBP localizations and availability, initiating further events that sequester more RNA and RBPs, thus disrupting RNA processing. The exact progression of events and outcomes are likely dependent on the cell type and overall transcriptional activity of the cell, which could potentially explain tissue specific disease phenotypes. The interaction between RNA with RBPs could lead to the formation of membrane-less phase separated compartments that could also impact nuclear processes and drive disease (Decker and Parker, 2012). Notably, various neurodegenerative diseases are linked to altered or defective RNA metabolism events, which in some instances have been shown to include the improper localization of RBPs (Volkening and J. Strong, 2011). In addition, viral pathogens have more recently been found to modulate exosome and RNA decay activities to both promote their own replication and evade host defense (Rialdi *et al.*, 2017; Guo *et al.*, 2018). Speculatively, the phenotypes and molecular mechanisms observed here in yeast involving the exosome and RBPs may underlie the mechanisms at play in these disease states and during viral infection. Given these links, and our evolving understanding of nuclear RNA processing, it is expected that future work will provide examples in both normal and pathological conditions where alterations in nuclear RNA homeostasis act to change cellular function and overall cellular physiology.

References

Adams, R. L., Terry, L. J., and Wente, S. R. (2014). Nucleoporin FG domains facilitate mRNP remodeling at the cytoplasmic face of the nuclear pore complex. Genetics *197*, 1213–1224.
Albert, B., Kos-Braun, I. C., Henras, A. K., Dez, C., Rueda, M. P., Zhang, X., Gadal, O., Kos, M., and Shore, D. (2019). A ribosome assembly stress response regulates transcription to maintain proteome homeostasis. Elife *8*.
Alcazar-Roman, A. R., Tran, E. J., Guo, S., Wente, S. R., and Alcázar-Román, A. R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. Nat Cell Biol *8*, 711–716.
Allison, L. A., Wong, J. K., Fitzpatrick, V. D., Moyle, M., and Ingles, C. J. (1988). The C-terminal domain of the largest subunit of RNA polymerase II of Saccharomyces cerevisiae, Drosophila melanogaster, and mammals: a conserved structure with an essential function. Mol. Cell. Biol. *8*, 321–329.
Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999a). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J. *18*, 5399–5410.
Allmang, C., Mitchell, P., Petfalski, E., and Tollervey, D. (2000). Degradation of ribosomal RNA precursors by the exosome. Nucleic Acids Res. *28*, 1684–1691.
Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999b). The yeast exosome and human PM-Scl are related complexes of 3’ right-arrow 5’ exonucleases. Genes Dev. *13*, 2148–2158.
Amberg, D. C., Fleischmann, M., Stagljar, I., Cole, C. N., and Aebi, M. (1993). Nuclear PRP20 protein is required for mRNA export. Embo J *12*, 233–241.
Amberg, D. C., Goldstein, A. L., and Cole, C. N. (1992). Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev. *6*, 1173–1189.
Amrani, N., Minet, M., Le Gouar, M., Lacroute, F., and Wyers, F. (1997). Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length in vitro. Mol Cell Biol *17*, 3694-701.
An, J. J. *et al.* (2008). Distinct Role of Long 3′ UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons. Cell *134*, 175–187.
Anderson, J. T., Wilson, S. M., Datar, K. V, and Swanson, M. S. (1993). NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. Mol Cell Biol *13*, 2730–2741.

Ares, M., Grate, L., and Pauling, M. H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA *5*, 1138–1139.

Arigo, J. T., Carroll, K. L., Ames, J. M., and Corden, J. L. (2006a). Regulation of Yeast NRD1 Expression by Premature Transcription Termination. Mol. Cell *21*, 641–651.

Arigo, J. T., Eyler, D. E., Carroll, K. L., and Corden, J. L. (2006b). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell *23*, 841–851.

Armache, J. P. *et al.* (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-Å resolution. Proc. Natl. Acad. Sci. U. S. A. *107*, 19748–19753.

Arndt, K. M., and Reines, D. (2015). Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. Annu. Rev. Biochem. *84*, 381–404.

Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J., and Cramer, P. (2014). Transcriptome Maps of mRNP Biogenesis Factors Define Pre-mRNA Recognition. Mol. Cell *55*, 745–757.

Balakin, A. G., Smith, L., and Fournier, M. J. (1996). The RNA World of the Nucleolus: Two Major Families of Small RNAs Defined by Different Box Elements with Related Functions. Cell *86*, 823–834.

Belgareh-Touzé, N., Léon, S., Erpapazoglou, Z., Stawiecka-Mirota, M., Urban-Grimal, D., and Haguenauer-Tsapis, R. (2008). Versatile role of the yeast ubiquitin ligase Rsp5p in intracellular trafficking. Biochem. Soc. Trans. *36*, 791–796.

Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K. A., Boeke, J. D., and Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae. Mol. Cell *30*, 248–258.

Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. Science *330*, 1203–1209.

Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R. H., and Engelke, D. R. (1998). Nucleolar localization of early tRNA processing. Genes Dev. *12*, 2463–2468.

Bhalla, N., Biggins, S., and Murray, A. W. (2002). Mutation of YCS4 , a Budding Yeast Condensin Subunit, Affects Mitotic and Nonmitotic Chromosome Behavior. Mol. Biol. Cell *13*, 632–645.

Black, D. L., Chabot, B., and Steitz, J. A. (1985). U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell *42*, 737–750.

Black, D. L., and Steitz, J. A. (1986). Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein. Cell *46*, 697–704.

Blattner, C. *et al.* (2011). Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth. Genes Dev. *25*, 2093–2105.

Blobel, G. (1985). Gene gating: a hypothesis. Proc. Natl. Acad. Sci. U. S. A. *82*, 8527–8529.

Boni, I. V., Lsaeva, D. M., Musychenko, M. L., and Tzareva, N. V. (1991). Ribosome-messenger recognition: MRNA target sites for ribosomal protein S1. Nucleic Acids Res. *19*, 155–162.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009). The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. Cell *139*, 547–559.

Bono, F., and Gehring, N. H. (2011). Assembly, disassembly and recycling: the dynamics of exon junction complexes. RNA Biol. *8*, 24–29.

Bose, T. *et al.* (2012). Cohesin proteins promote ribosomal RNA production and protein translation in yeast and human cells. PLoS Genet. *8*, e1002749.

Breslow, D. K., Cameron, D. M., Collins, S. R., Schuldiner, M., Stewart-Ornstein, J., Newman, H. W., Braun, S., Madhani, H. D., Krogan, N. J., and Weissman, J. S. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. Nat. Methods *5*, 711–718.

Bresson, S., and Tollervey, D. (2018). Surveillance-ready transcription: Nuclear RNA decay as a default fate. Open Biol. *8*.

Briggs, M. W., Burkard, K. T., and Butler, J. S. (1998). 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. *273*, 13255–13263.

Brockmann, C. *et al.* (2012). Structural basis for polyadenosine-RNA binding by Nab2 Zn fingers and its function in mRNA nuclear export. Structure *20*, 1007–1018.

Brodsky, A. S., and Silver, P. A. (2000). Pre-mRNA processing factors are required for nuclear export. RNA *6*, 1737–1749.

Brogna, S., and Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nat. Struct. Mol. Biol. *16*, 107–113.

Brown, A., and Shao, S. (2018). Ribosomes and cryo-EM: a duet. Curr. Opin. Struct. Biol. *52*, 1–7.

Brune, C., Munchel, S. E., Fischer, N., Podtelejnikov, A. V, and Weis, K. (2005). Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. RNA *11*, 517–531.

de Bruyn Kops, A., and Guthrie, C. (2001). An essential nuclear envelope integral membrane protein, Brr6p, required for nuclear transport. EMBO J. *20*, 4183–4193.

Burkard, K. T., and Butler, J. S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol Cell Biol *20*, 604-16.

Callahan, K. P., and Butler, J. S. (2008). Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. Nucleic Acids Res. *36*, 6645–6655.

Camblong, J., Beyrouthy, N., Guffanti, E., Schlaepfer, G., Steinmetz, L. M., and Stutz, F. (2009). Trans-acting antisense RNAs mediate transcriptional gene cosuppression in S. cerevisiae. Genes Dev. *23*, 1534–1545.

Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., and Stutz, F. (2007). Antisense RNA Stabilization Induces Transcriptional Gene Silencing via Histone Deacetylation in S. cerevisiae. Cell *131*, 706–717.

Canavan, R., and Bond, U. (2007). Deletion of the nuclear exosome component RRP6 leads to continued accumulation of the histone mRNA HTB1 in S-phase of the cell cycle in Saccharomyces cerevisiae. Nucleic Acids Res. *35*, 6268–6279.

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature *515*, 143–146.

Carneiro, T., Carvalho, C., Braga, J., Rino, J., Milligan, L., Tollervey, D., and Carmo-Fonseca, M. (2007). Depletion of the yeast nuclear exosome subunit Rrp6 results in accumulation of polyadenylated RNAs in a discrete domain within the nucleolus. Mol. Cell. Biol. *27*, 4157–4165.

Carroll, K. L., Ghirlando, R., Ames, J. M., and Corden, J. L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. RNA *13*, 361–373.

Carrozza, M. J. *et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Casolari, J. M., Brown, C. R., Komili, S., West, J., Hieronymus, H., and Silver, P. A. (2004).

Genome-Wide Localization of the Nuclear Transport Machinery Couples Transcriptional Status and Nuclear Organization. Cell *117*, 427–439.

Castelnuovo, M., Rahman, S., Guffanti, E., Infantino, V., Stutz, F., and Zenklusen, D. (2013). Bimodal expression of PHO84 is modulated by early termination of antisense transcription. Nat. Struct. Mol. Biol. *20*, 851–858.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chazal, P. E., Daguenet, E., Wendling, C., Ulryck, N., Tomasetto, C., Sargueil, B., and Le Hir, H. (2013). EJC core component MLN51 interacts with eIF3 and activates translation. Proc. Natl. Acad. Sci. U. S. A. *110*, 5903–5908.

Chen, J., and Moore, C. (1992). Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. Mol. Cell. Biol. *12*, 3470–3481.

Chen, W., Bucaria, J., Band, D. A., Sutton, A., and Sternglanz, R. (2003). Enp1, a yeast protein associated with U3 and U14 snoRNAs, is required for pre-rRNA processing and 40S subunit synthesis. Nucleic Acids Res. *31*, 690–699.

Chlebowski, A., Lubas, M., Jensen, T. H., and Dziembowski, A. (2013). RNA decay machines: the exosome. Biochim. Biophys. Acta *1829*, 552–560.

Cho, E. J., Rodriguez, C. R., Takagi, T., and Buratowski, S. (1998). Allosteric interactions between capping enzyme subunits and the RNA polymerase II. Carboxy-terminal domain. Genes Dev. *12*, 3482–3487.

Choque, E., Schneider, C., Gadal, O., and Dez, C. (2018). Turnover of aberrant pre-40S pre-ribosomal particles is initiated by a novel endonucleolytic decay pathway. Nucleic Acids Res. *46*, 4699–4714.

Churchman, L. S., and Weissman, J. S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Churchman, L. S., and Weissman, J. S. (2012). Native elongating transcript sequencing (NET-seq). Curr. Protoc. Mol. Biol. *1*, 14.4.1-14.4.17.

Ciganda, M., and Williams, N. (2011). Eukaryotic 5S rRNA biogenesis. Wiley Interdiscip. Rev. RNA *2*, 523–533.

Cole, C. N., Heath, C. V, Hodge, C. A., Hammell, C. M., and Amberg, D. C. (2002). Analysis of

RNA export. Methods Enzymol. *351*, 568–587.

Cole, S. E., LaRiviere, F. J., Merrikh, C. N., and Moore, M. J. (2009). A Convergence of rRNA and mRNA Quality Control Pathways Revealed by Mechanistic Analysis of Nonfunctional rRNA Decay. Mol. Cell *34*, 440–450.

Collart, M. A., and Oliviero, S. (1993). Preparation of Yeast RNA. Curr. Protoc. Mol. Biol. *23*, 13.12.1-13.12.5.

Conaway, R. C., Bradsher, J. N., and Conaway, J. W. (1992). Mechanism of assembly of the RNA polymerase II preinitiation complex. Evidence for a functional interaction between the carboxyl-terminal domain of the largest subunit of RNA polymerase II and a high molecular mass form of the TATA factor. J. Biol. Chem. *267*, 8464–8467.

Conrad, N. K., Wilson, S. M., Steinmetz, E. J., Patturajan, M., Brow, D. A., Swanson, M. S., and Corden, J. L. (2000). A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. Genetics *154*, 557–571.

Coppola, J. A., Field, A. S., and Luse, D. S. (1983). Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. Proc. Natl. Acad. Sci. U. S. A. *80*, 1251–1255.

D'Amours, D., Stegmeier, F., and Amon, A. (2004). Cdc14 and Condensin Control the Dissolution of Cohesin-Independent Chromosome Linkages at Repeated DNA. Cell *117*, 455–469.

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C. J., Bofkin, L., Jones, T., Davis, R. W., and Steinmetz, L. M. (2006). A high-resolution map of transcription in the yeast genome. Proc. Natl. Acad. Sci. U. S. A. *103*, 5320–5325.

Davis, C. A., and Ares, M. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *103*, 3262–3267.

Decker, C. J., and Parker, R. (2012). P-bodies and stress granules: Possible roles in the control of translation and mRNA degradation. Cold Spring Harb. Perspect. Biol. *4*, a012286–a012286.

Dez, C., Houseley, J., and Tollervey, D. (2006). Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of Saccharomyces cerevisiae. EMBO J. *25*, 1534–1546.

Dheur, S., Nykamp, K. R., Viphakone, N., Swanson, M. S., and Minvielle-Sebastia, L. (2005). Yeast mRNA poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-

dependent poly(A) nuclease activity. J. Biol. Chem. *280*, 24532–24538.
Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase 11 in a soluble extract from isolated mammalian nuclei. Nucleic. Acids. Res. *11*, 1475–1489.
Van Dijk, E. L. *et al.* (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114–119.
Dockendorff, T., Heath, C., Goldstein, A., Snay, C., and Cole, C. (1997). C-terminal truncations of the yeast nucleoporin Nup145p produce a rapid temperature-conditional mRNA export defect and alterations to nuclear structure [published erratum appears in Mol Cell Biol 1997 Apr;17(4):2347-50]. Mol. Cell. Biol. *17*, 906–920.
Dowen, J. M., Bilodeau, S., Orlando, D. A., Hübner, M. R., Abraham, B. J., Spector, D. L., and Young, R. A. (2013). Multiple Structural Maintenance of Chromosome Complexes at Transcriptional Regulatory Elements. Stem Cell Reports *1*, 371–378.
Dowen, J. M., and Young, R. A. (2014). SMC complexes link gene expression and genome architecture. Curr. Opin. Genet. Dev. *25*, 131–137.
Doye, V. R., Wepf, R., and Hurt, E. C. (1994). A novel nuclear pore protein Nup133p with distinct roles in poly(A)+RNA transport and nuclear pore distribution. EMBO J *13*, 6062–6075.
Dragon, F. *et al.* (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature *417*, 967–970.
Du, T.-G., Jellbauer, S., Müller, M., Schmid, M., Niessing, D., and Jansen, R.-P. (2008). Nuclear transit of the RNA-binding protein She2 is required for translational control of localized ASH1 mRNA. EMBO Rep. *9*, 781–787.
Duncan, K., Umen, J. G., and Guthrie, C. (2000). A putative ubiquitin ligase required for efficient mRNA export differentially affects hnRNP transport. Curr Biol *10*, 687–696.
Dürrbaum, M., and Storchová, Z. (2016). Effects of aneuploidy on gene expression: implications for cancer. FEBS J. *283*, 791–802.
Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat. Struct. Mol. Biol. *14*, 15–22.
Dziembowski, A., Ventura, A.-P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprévote, O., and Séraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. EMBO J. *23*, 4847–4856.

Eberle, A. B., and Visa, N. (2014). Quality control of mRNP biogenesis: networking at the transcription site. Semin. Cell Dev. Biol. *32*, 37–46.

Egecioglu, D. E., Henras, A. K., and Chanfreau, G. F. (2006). Contributions of Trf4p- And Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA *12*, 26–32.

Elela, S. A., Igel, H., and Ares, M. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell *85*, 115–124.

Fabre, A. *et al.* (2012). SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome. Am. J. Hum. Genet. *90*, 689–692.

Fabre, E., Boelens, W. C., Wimmer, C., Mattaj, I. W., and Hurt, E. C. (1994). Nup145p is required for nuclear export of mRNA and binds homopolymeric RNA in vitro via a novel conserved motif. Cell *78*, 275–289.

Falk, S., Bonneau, F., Ebert, J., Kögel, A., and Conti, E. (2017a). Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase. Cell Rep. *20*, 2279–2286.

Falk, S., Tants, J. N., Basquin, J., Thoms, M., Hurt, E., Sattler, M., and Conti, E. (2017b). Structural insights into the interaction of the nuclear exosome helicase Mtr4 with the preribosomal protein Nop53. RNA *23*, 1780–1787.

Fasken, M. B., and Corbett, A. H. (2009). Mechanisms of nuclear mRNA quality control. RNA Biol. *6*, 237–241.

Fasken, M. B., Stewart, M., and Corbett, A. H. (2008). Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, mlp1, in mRNA export. J. Biol. Chem. *283*, 27130–27143.

Fatica, A., and Tollervey, D. (2002). Making ribosomes. Curr Opin Cell Biol *14*, 313-8.

Fica, S. M., Tuttle, N., Novak, T., Li, N. S., Lu, J., Koodathingal, P., Dai, Q., Staley, J. P., and Piccirilli, J. A. (2013). RNA catalyses nuclear pre-mRNA splicing. Nature *503*, 229–234.

Fortes, P., Inada, T., Preiss, T., Hentze, M. W., Mattaj, I. W., and Sachs, A. B. (2000). The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol Cell *6*, 191–196.

Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., and Mattaj, I. W. (1999). Genetic and physical interactions involving the yeast nuclear cap-binding

complex. Mol Cell Biol *19*, 6543–6553.

Fox, M. J., Gao, H., Smith-Kinnaman, W. R., Liu, Y., and Mosley, A. L. (2015). The Exosome Component Rrp6 Is Required for RNA Polymerase II Termination at Specific Targets of the Nrd1-Nab3 Pathway. PLoS Genet.

Freeman, L. (2000). The Condensin Complex Governs Chromosome Condensation and Mitotic Transmission of rDNA. J. Cell Biol. *149*, 811–824.

Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell *116*, 63–73.

Gautier, T., Bergès, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol. Cell. Biol. *17*, 7088–7098.

Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W., and Kulozik, A. E. (2003). Y14 and hUpf3b form an NMD-activating complex. Mol. Cell *11*, 939–949.

Gelperin, D. M. *et al.* (2005). Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. Genes Dev. *19*, 2816–2826.

Gewartowski, K., Cuéllar, J., Dziembowski, A., and Valpuesta, J. M. (2012). The yeast THO complex forms a 5-subunit assembly that directly interacts with active chromatin. Bioarchitecture *2*.

Ghosh, S., and Jacobson, A. (2010). RNA decay modulates gene expression and controls its fidelity. Wiley Interdiscip. Rev. RNA *1*, 351–361.

Gietz, D., St, J. A., Woods, R. A., and Schiestl, R. H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res *20*, 1425.

Gilbert, W., and Guthrie, C. (2004). The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. Mol. Cell *13*, 201–212.

Gingras, A.-C., Raught, B., and Sonenberg, N. (1999). eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation. Annu. Rev. Biochem. *68*, 913–963.

Giunta, M., Edvardson, S., Xu, Y., Schuelke, M., Gomez-Duran, A., Boczonadi, V., Elpeleg, O., Müller, J. S., and Horvath, R. (2016). Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. Hum. Mol. Genet. *25*, 2985–2996.

Glynn, E. F., Megee, P. C., Yu, H.-G., Mistrot, C., Unal, E., Koshland, D. E., DeRisi, J. L., and Gerton, J. L. (2004). Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS Biol. *2*, E259.

Goldstein, A. L., Snay, C. A., Heath, C. V, and Cole, C. N. (1996). Pleiotropic nuclear defects associated with a conditional allele of the novel nucleoporin Rat9p/Nup85p. Mol. Biol. Cell *7*, 917–934.

González-Aguilera, C., Tous, C., Babiano, R., de la Cruz, J., Luna, R., and Aguilera, A. (2011). Nab2 functions in the metabolism of RNA driven by polymerases II and III. Mol. Biol. Cell *22*, 2729–2740.

Gordon, D. J., Resio, B., and Pellman, D. (2012). Causes and consequences of aneuploidy in cancer. Nat. Rev. Genet. *13*, 189–203.

Gorsch, L. C., Dockendorff, T. C., and Cole, C. N. (1995). A conditional allele of the novel repeat-containing yeast nucleoporin RAT7/NUP159 causes both rapid cessation of mRNA export and reversible clustering of nuclear pore complexes. J Cell Biol *129*, 939–955.

Greber, B. J., Boehringer, D., Montellese, C., and Ban, N. (2012). Cryo-EM structures of Arx1 and maturation factors Rei1 and Jjj1 bound to the 60S ribosomal subunit. Nat. Struct. Mol. Biol. *19*, 1228–1233.

Green, D. M., Marfatia, K. A., Crafton, E. B., Zhang, X., Cheng, X., and Corbett, A. H. (2002). Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. J. Biol. Chem. *277*, 7752–7760.

Grünwald, D., and Singer, R. H. (2010). In vivo imaging of labelled endogenous β-actin mRNA during nucleocytoplasmic transport. Nature *467*, 604–607.

Grzechnik, P., and Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Mol. Cell *32*, 247–258.

Gudipati, R. K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L. M., Jacquier, A., and Libri, D. (2012). Extensive degradation of RNA precursors by the exosome in wild-type cells. Mol. Cell *48*, 409–421.

Guo, L., Louis, I. V. S., and Bohjanen, P. R. (2018). Viral manipulation of host mRNA decay. Future Virol. *13*, 211–223.

Gupta, R., Kus, B., Fladd, C., Wasmuth, J., Tonikian, R., Sidhu, S., Krogan, N. J., Parkinson, J., and Rotin, D. (2007). Ubiquitination screen using protein microarrays for comprehensive

identification of Rsp5 substrates in yeast. Mol. Syst. Biol. *3*, 116.

Gwizdek, C., Hobeika, M., Kus, B., Ossareh-Nazari, B., Dargemont, C., and Rodriguez, M. S. (2005). The mRNA nuclear export factor Hpr1 is regulated by Rsp5-mediated ubiquitylation. J. Biol. Chem. *280*, 13401–13405.

Hackmann, A., Wu, H., Schneider, U.-M., Meyer, K., Jung, K., and Krebber, H. (2014). Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. Nat. Commun. *5*, 3123.

Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297–301.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Genes Dev. *22*, 1069–1081.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The yeast ski complex: Crystal structure and rna channeling to the exosome complex. Cell *154*, 814–826.

Hammell, C. M., Gross, S., Zenklusen, D., Heath, C. V, Stutz, F., Moore, C., and Cole, C. N. (2002). Coupling of termination, 3' processing, and mRNA export. Mol Cell Biol *22*, 6441-57.

Han, J., and van Hoof, A. (2016). The RNA Exosome Channeling and Direct Access Conformations Have Distinct In Vivo Functions. Cell Rep. *16*, 3348–3358.

Hardwick, S. W., and Luisi, B. F. (2013). Rarely at rest. RNA Biol. *10*, 56–70.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., Chikashige, Y., Hiraoka, Y., Yamashita, A., and Yamamoto, M. (2006). Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature *442*, 45–50.

Harris, B. *et al.* (2014). Cohesion promotes nucleolar structure and function. Mol. Biol. Cell *25*, 337–346.

Hartley, J. L. *et al.* (2010). Mutations in TTC37 Cause Trichohepatoenteric Syndrome (Phenotypic Diarrhea of Infancy). Gastroenterology *138*, 2388-2398.e2.

Hazelbaker, D. Z., Marquardt, S., Wlotzka, W., and Buratowski, S. (2013). Kinetic Competition between RNA Polymerase II and Sen1-Dependent Transcription Termination. Mol. Cell *49*, 55–66.

He, H., Zhang, S., Wang, D., Hochwagen, A., and Li, F. (2016). Condensin Promotes Position Effects within Tandem DNA Repeats via the RITS Complex. Cell Rep. *14*, 1018–1024.

Hector, R. E., Nykamp, K. R., Dheur, S., Anderson, J. T., Non, P. J., Urbinati, C. R., Wilson, S.

M., Minvielle-Sebastia, L., and Swanson, M. S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. Embo J *21*, 1800-10.
Henras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A., and Gleizes, P. E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip. Rev. RNA *6*, 225–242.
Henry, Y., Wood, H., Morrissey, J. P., Petfalski, E., Kearsey, S., and Tollervey, D. (1994). The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. EMBO J. *13*, 2452–2463.
Heo, D., Yoo, I., Kong, J., Lidschreiber, M., Mayer, A., Choi, B.-Y., Hahn, Y., Cramer, P., Buratowski, S., and Kim, M. (2013). The RNA polymerase II C-terminal domain-interacting domain of yeast Nrd1 contributes to the choice of termination pathway and couples to RNA processing by the nuclear exosome. J. Biol. Chem. *288*, 36676–36690.
Hernandez-Verdun, D. (2006). Nucleolus: From structure to dynamics. Histochem. Cell Biol. *125*, 127–137.
Hieronymus, H., and Silver, P. A. (2003). Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery. Nat. Genet. *33*, 155–161.
Hieronymus, H., Yu, M. C., and Silver, P. A. (2004). Genome-wide mRNA surveillance is coupled to mRNA export. Genes Dev. *18*, 2652–2662.
Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T. H. (2001). Quality control of mRNA 3'-end processing is linked to the nuclear exosome. Nature *413*, 538-42.
Hilleren, P., and Parker, R. (2001). Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. Rna *7*, 753-64.
Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M. J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. Embo J *20*, 4987-97.
Le Hir, H., Moore, M. J., and Maquat, L. E. (2000). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev *14*, 1098–1108.
Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. Nat. Rev. Mol. Cell Biol. *7*, 311–322.
Hocquet, C. *et al.* (2018). Condensin controls cellular RNA levels through the accurate

segregation of chromosomes instead of directly regulating transcription. Elife *7*.
Hodge, C. A., Choudhary, V., Wolyniak, M. J., Scarcelli, J. J., Schneiter, R., and Cole, C. N. (2010). Integral membrane proteins Brr6 and Apq12 link assembly of the nuclear pore complex to lipid homeostasis in the endoplasmic reticulum. J. Cell Sci. *123*, 141–151.
Hodge, C. a, Tran, E. J., Noble, K. N., Alcazar-Roman, A. R., Ben-Yishay, R., Scarcelli, J. J., Folkmann, A. W., Shav-Tal, Y., Wente, S. R., and Cole, C. N. (2011). The Dbp5 cycle at the nuclear pore complex during mRNA export I: dbp5 mutants with defects in RNA binding and ATP hydrolysis define key steps for Nup159 and Gle1. Genes Dev. *25*, 1052–1064.
Holland, A. J., and Cleveland, D. W. (2012). Losing balance: the origin and impact of aneuploidy in cancer. EMBO Rep. *13*, 501–514.
Hongay, C. F., Grisafi, P. L., Galitski, T., and Fink, G. R. (2006). Antisense Transcription Controls Cell Fate in Saccharomyces cerevisiae. Cell *127*, 735–745.
van Hoof, A., Lennertz, P., and Parker, R. (2000a). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol. Cell. Biol. *20*, 441–452.
van Hoof, A., Staples, R. R., Baker, R. E., and Parker, R. (2000b). Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. Mol Cell Biol *20*, 8230-43.
Hopper, A. K., and Huang, H.-Y. (2015). Quality Control Pathways for Nucleus-Encoded Eukaryotic tRNA Biosynthesis and Subcellular Trafficking. Mol. Cell. Biol. *35*, 2052–2058.
Hoque, M., Ji, Z., Zheng, D., Luo, W., Li, W., You, B., Park, J. Y., Yehia, G., and Tian, B. (2013). Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. Nat. Methods *10*, 133–139.
Houalla, R., Devaux, F., Fatica, A., Kufel, J., Barrass, D., Torchet, C., and Tollervey, D. (2006). Microarray detection of novel nuclear RNA substrates for the exosome. Yeast *23*, 439–454.
Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA Modulates Histone Modification and mRNA Induction in the Yeast GAL Gene Cluster. Mol. Cell *32*, 685–695.
Houseley, J., and Tollervey, D. (2006). Yeast Trf5p is a nuclear poly(A) polymerase. EMBO Rep. *7*, 205–211.
Houseley, J., and Tollervey, D. (2009). The many pathways of RNA degradation. Cell *136*, 763–776.

Hsin, J. P., and Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119–2137.

Huh, W.-K. K., Falvo, J. V, Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., and O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686–691.

Hurt, E., Strasser, K., Segref, A., Bailer, S., Schlaich, N., Presutti, C., Tollervey, D., and Jansen, R. (2000). Mex67p mediates nuclear export of a variety of RNA polymerase II transcripts. J Biol Chem *275*, 8361–8368.

Ideue, T., Azad, A. K., Yoshida, J., Matsusaka, T., Yanagida, M., Ohshima, Y., and Tani, T. (2004). The nucleolus is involved in mRNA export from the nucleus in fission yeast. J. Cell Sci. *117*, 2887–2895.

Iglesias, N., Tutucci, E., Gwizdek, C., Vinciguerra, P., Von Dach, E., Corbett, A. H., Dargemont, C., and Stutz, F. (2010). Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export. Genes Dev. *24*, 1927–1938.

Itoh, N., Yamada, H., Kaziro, Y., and Mizumoto, K. (1987). Messenger RNA guanylyltransferase from Saccharomyces cerevisiae. Large scale purification, subunit functions, and subcellular localization. J. Biol. Chem. *262*, 1989–1995.

Iwasaki, O., Tanaka, A., Tanizawa, H., Grewal, S. I. S., and Noma, K. (2010). Centromeric Localization of Dispersed Pol III Genes in Fission Yeast. Mol. Biol. Cell *21*, 254–265.

Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1995). A cap-binding protein complex mediating U snRNA export. Nature *376*, 709–712.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657–668.

Izawa, S. (2010). Ethanol stress response in the mRNA flux of saccharomyces cerevisiae. Biosci. Biotechnol. Biochem. *74*, 7–12.

Jackson, R. N., Klauer, A. A., Hintze, B. J., Robinson, H., Van Hoof, A., and Johnson, S. J. (2010). The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. EMBO J. *29*, 2205–2216.

Jacquier, A. (2009). The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. Nat. Rev. Genet. *10*, 833–844.

Janicke, A., Vancuylenberg, J., Boag, P. R., Traven, A., and Beilharz, T. H. (2012). ePAT: A simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. RNA *18*, 1289–1295.

Jensen, T. H., Boulay, J., Rosbash, M., and Libri, D. (2001a). The DECD box putative ATPase Sub2p is an early mRNA export factor. Curr Biol *11*, 1711-5.

Jensen, T. H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. Mol. Cell *52*, 473–484.

Jensen, T. H., Patricio, K., McCarthy, T., and Rosbash, M. (2001b). A block to mRNA nuclear export in S. cerevisiae leads to hyperadenylation of transcripts that accumulate at the site of transcription. Mol Cell *7*, 887-98.

Jiao, X., Xiang, S., Oh, C., Martin, C. E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5′-end capping. Nature *467*, 608–611.

Johnson, S. A., Cubberley, G., and Bentley, D. L. (2009). Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. Mol. Cell *33*, 215–226.

Jorjani, H., Kehr, S., Jedlinski, D. J., Gumienny, R., Hertel, J., Stadler, P. F., Zavolan, M., and Gruber, A. R. (2016). An updated human snoRNAome. Nucleic Acids Res. *44*, 5068–5082.

Kadaba, S., Wang, X., and Anderson, J. T. (2006). Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. RNA *12*, 508–521.

Kadowaki, T., Chen, S., Hitomi, M., Jacobs, E., Kumagai, C., Liang, S., Schneiter, R., Singleton, D., Wisniewska, J., and Tartakoff, A. M. (1994a). Isolation and characterization of Saccharomyces cerevisiae mRNA transport-defective (mtr) mutants. J. Cell Biol. *126*, 649–659.

Kadowaki, T., Chen, S., Hitomi, M., Jacobs, E., Kumagai, C., Liang, S., Schneiter, R., Singleton, D., Wisniewska, J., and Tartakoff, A. M. (1994b). Isolation and characterization of Saccharomyces cerevisiae mRNA transport-defective (mtr) mutants [published erratum appears in J Cell Biol 1994 Sep;126(6):1627]. J Cell Biol *126*, 649–659.

Kadowaki, T., Hitomi, M., Chen, S., and Tartakoff, A. M. (1994c). Nuclear mRNA accumulation causes nucleolar fragmentation in yeast mtr2 mutant. Mol. Biol. Cell *5*, 1253–1263.

Kadowaki, T., Zhao, Y., and Tartakoff, A. M. (1992). A conditional yeast mutant deficient in mRNA transport from nucleus to cytoplasm. Proc. Natl. Acad. Sci. *89*, 2312–2316.

Kalam Azad, A., Ideue, T., Ohshima, Y., and Tani, T. (2003). A mutation in the gene involved in sister chromatid separation causes a defect in nuclear mRNA export in fission yeast. Biochem Biophys Res Commun *310*, 176–181.
Kaliszewski, P., and Zoładek, T. (2008). The role of Rsp5 ubiquitin ligase in regulation of diverse processes in yeast cells. Acta Biochim. Pol. *55*, 649–662.
Kallehauge, T. B., Robert, M.-C., Bertrand, E., and Jensen, T. H. (2012). Nuclear retention prevents premature cytoplasmic appearance of mRNA. Mol. Cell *48*, 145–152.
Kaplan, C. D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096–1099.
Kapranov, P. *et al.* (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.
Katahira, J. *et al.* (1999). The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. Embo J *18*, 2593–2609.
Kessler, M. M., Henry, M. F., Shen, E., Zhao, J., Gross, S., Silver, P. A., and Moore, C. L. (1997). Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. Genes Dev. *11*, 2545–2556.
Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.
Kim, J. H., Zhang, T., Wong, N. C., Davidson, N., Maksimovic, J., Oshlack, A., Earnshaw, W. C., Kalitsis, P., and Hudson, D. F. (2013). Condensin i associates with structural and gene regulatory regions in vertebrate chromosomes. Nat. Commun. *4*, 2537.
Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517–522.
Kim, M., Vasiljeva, L., Rando, O. J., Zhelkovsky, A., Moore, C., and Buratowski, S. (2006). Distinct pathways for snoRNA and mRNA termination. Mol. Cell *24*, 723–734.
Kiss-László, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J. *17*, 797–807.
Kiss, T. (2002). Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions. Cell *109*, 145–148.

Klauer, A. A., and van Hoof, A. (2013). Genetic interactions suggest multiple distinct roles of the arch and core helicase domains of Mtr4 in Rrp6 and exosome function. Nucleic Acids Res. *41*, 533–541.

Klinge, S., and Woolford, J. L. (2019). Ribosome assembly coming into focus. Nat. Rev. Mol. Cell Biol. *20*, 116–131.

Koepp, D. M., Wong, D. H., Corbett, A. H., and Silver, P. A. (1996). Dynamic localization of the nuclear import receptor and its interactions with transport factors. J. Cell Biol. *133*, 1163–1176.

Köhler, A., Hurt, E., Kohler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol *8*, 761–773.

Konarska, M. M., Grabowski, P. J., Padgett, R. A., and Sharp, P. A. (1985). Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature *313*, 552–557.

Konarska, M. M., and Sharp, P. A. (1987). Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell *49*, 763–774.

Koš, M., and Tollervey, D. (2010). Yeast Pre-rRNA Processing and Modification Occur Cotranscriptionally. Mol. Cell *37*, 809–820.

Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016). Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. Mol. Cell *63*, 125–134.

Krantz, I. D. *et al.* (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat. Genet. *36*, 631–635.

Kress, T. L., Krogan, N. J., and Guthrie, C. (2008). A Single SR-like Protein, Npl3, Promotes Pre-mRNA Splicing in Budding Yeast. Mol. Cell *32*, 727–734.

Kressler, D., Hurt, E., and Baßler, J. (2017). A Puzzle of Life: Crafting Ribosomal Subunits. Trends Biochem. Sci. *42*, 640–654.

Kuai, L., Fang, F., Butler, J. S., and Sherman, F. (2004). Polyadenylation of rRNA in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *101*, 8581–8586.

Kufel, J., and Grzechnik, P. (2019). Small Nucleolar RNAs Tell a Different Tale. Trends Genet. *35*, 104–117.

Kyburz, A., Sadowski, M., Dichtl, B., and Keller, W. (2003). The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3′-end formation. Nucleic Acids Res. *31*, 3936–3945.

de la Cruz, J., Kressler, D., Tollervey, D., and Linder, P. (1998). Dob1p (Mtr4p) is a putative

ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. EMBO J. *17*, 1128–1140.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713–724.

Lardenois, A., Liu, Y., Walther, T., Chalmel, F., Evrard, B., Granovskaia, M., Chu, A., Davis, R. W., Steinmetz, L. M., and Primig, M. (2011). Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. Proc. Natl. Acad. Sci. U. S. A. *108*, 1058–1063.

Lebreton, A., Tomecki, R., Dziembowski, A., and Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993–996.

Lee, M. S., Henry, M., and Silver, P. A. (1996). A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. Genes Dev *10*, 1233–1246.

Lemay, J.-F., D'Amours, A., Lemieux, C., Lackner, D. H., St-Sauveur, V. G., Bähler, J., and Bachand, F. (2010). The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. Mol. Cell *37*, 34–45.

Leporé, N., and Lafontaine, D. L. J. (2011). A functional interface at the rDNA connects rRNA synthesis, pre-rRNA processing and nucleolar surveillance in budding yeast. PLoS One *6*, e24962.

Li, Z. *et al.* (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat. Biotechnol. *29*, 361–367.

Liang, W. Q., and Fournier, M. J. (1995). U14 base-pairs with 18S rRNA: A novel snoRNA interaction required for rRNA processing. Genes Dev. *9*, 2433–2443.

Liao, Y., Smyth, G. K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Libri, D., Dower, K., Boulay, J., Thomsen, R., Rosbash, M., and Jensen, T. H. (2002). Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. Mol Cell Biol *22*, 8254-66.

Lima, S. A., Chipman, L. B., Nicholson, A. L., Chen, Y. H., Yee, B. A., Yeo, G. W., Coller, J., and Pasquinelli, A. E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat. Struct. Mol. Biol. *24*, 1057–1063.

Lin, C. H., MacGurn, J. A., Chu, T., Stefan, C. J., and Emr, S. D. (2008). Arrestin-related ubiquitin-ligase adaptors regulate endocytosis and protein turnover at the cell surface. Cell *135*, 714–725.

Lindahl, L., Bommankanti, A., Li, X., Hayden, L., Jones, A., Khan, M., Tolulope, O. N. I., and Zengel, J. M. (2009). RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. RNA *15*, 1407–1416.

Lingner, J., Kellermann, J., and Keller, W. (1991). Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature *354*, 496–498.

Liu, J. J., Bratkowski, M. A., Liu, X., Niu, C. Y., Ke, A., and Wang, H. W. (2014). Visualization of distinct substrate-recruitment pathways in the yeast exosome by em. Nat. Struct. Mol. Biol. *21*, 95–102.

Liu, Q., Greimann, J. C., and Lima, C. D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223–1237.

Loewith, R., and Hall, M. N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics *189*, 1177–1201.

Lone, M. A., Atkinson, A. E., Hodge, C. A., Cottier, S., Martínez-Montañés, F., Maithel, S., Mène-Saffrané, L., Cole, C. N., and Schneiter, R. (2015). Yeast Integral Membrane Proteins Apq12, Brl1, and Brr6 Form a Complex Important for Regulation of Membrane Homeostasis and Nuclear Pore Complex Biogenesis. Eukaryot. Cell *14*, 1217–1227.

Longtine, M. S., McKenzie 3rd, A., Demarini, D. J., Shah, N. G., Wach, A., Brachat, A., Philippsen, P., Pringle, J. R., McKenzie 3, A., and McKenzie, A. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953–961.

Lopez, P. J., and Séraphin, B. (1999). Genomic-scale quantitative analysis of yeast pre-mRNA splicing: implications for splice-site recognition. RNA *5*, 1135–1137.

Lorentzen, E., Dziembowski, A., Lindner, D., Seraphin, B., and Conti, E. (2007). RNA channelling by the archaeal exosome. EMBO Rep. *8*, 470–476.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc. Natl. Acad.

Sci. U. S. A. *88*, 10004–10008.
Lund, M. K., and Guthrie, C. (2005). The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Mol. Cell *20*, 645–651.
Luo, W., Johnson, A. W., and Bentley, D. L. (2006). The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: Implications for a unified allosteric-torpedo model. Genes Dev. *20*, 954–965.
Lustig, A. J., Lin, R. J., and Abelson, J. (1986). The yeast RNA gene products are essential for mRNA splicing in vitro. Cell *47*, 953–963.
Lykke-Andersen, S., Brodersen, D. E., and Jensen, T. H. (2009). Origins and activities of the eukaryotic exosome. J. Cell Sci. *122*, 1487–1494.
Ma, X. M., Yoon, S. O., Richardson, C. J., Jülich, K., and Blenis, J. (2008). SKAR Links Pre-mRNA Splicing to mTOR/S6K1-Mediated Enhanced Translation Efficiency of Spliced mRNAs. Cell *133*, 303–313.
Machín, F., Torres-Rosell, J., Jarmuz, A., and Aragón, L. (2005). Spindle-independent condensation-mediated segregation of yeast ribosomal DNA in late anaphase. J. Cell Biol. *168*, 209–219.
Makino, D. L., Baumgärtner, M., and Conti, E. (2013). Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature *495*, 70–75.
Makino, D. L., Schuch, B., Stegmann, E., Baumgärtner, M., Basquin, C., and Conti, E. (2015). RNA degradation paths in a 12-subunit nuclear exosome complex. Nature *524*, 54–58.
Malet, H. *et al.* (2010). RNA channelling by the eukaryotic exosome. EMBO Rep. *11*, 936–942.
Mandel, C. R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3′-end processing. Cell. Mol. Life Sci. *65*, 1099–1122.
Mansour, F. H., and Pestov, D. G. (2013). Separation of long RNA by agarose-formaldehyde gel electrophoresis. Anal. Biochem. *441*, 18–20.
Mao, X., Schwer, B., and Shuman, S. (1995). Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. Mol. Cell. Biol. *15*, 4167–4174.
Marquardt, S., Hazelbaker, D. Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3’ extensions of yeast non-coding RNA species. Transcription *2*, 145–154.
Martens, J. A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *429*, 571–574.

Matera, A. G., Terns, R. M., and Terns, M. P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat. Rev. Mol. Cell Biol. *8*, 209–220.

Meinel, D. M., Burkert-Kautzsch, C., Kieser, A., O'Duibhir, E., Siebert, M., Mayer, A., Cramer, P., Söding, J., Holstege, F. C. P., and Sträßer, K. (2013). Recruitment of TREX to the Transcription Machinery by Its Direct Binding to the Phospho-CTD of RNA Polymerase II. PLoS Genet. *9*, e1003914.

Milligan, L., Decourty, L., Saveanu, C., Rappsilber, J., Ceulemans, H., Jacquier, A., and Tollervey, D. (2008). A Yeast Exosome Cofactor, Mpp6, Functions in RNA Surveillance and in the Degradation of Noncoding RNA Transcripts. Mol. Cell. Biol. *28*, 5446–5457.

Minvielle-Sebastia, L., Preker, P. J., Wiederkehr, T., Strahm, Y., and Keller, W. (1997). The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. Proc Natl Acad Sci U S A *94*, 7897-902.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M., and Tollervey, D. (2003). Rrp47p Is an Exosome-Associated Protein Required for the 3' Processing of Stable RNAs. Mol. Cell. Biol. *23*, 6982–6992.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3′→5′ Exoribonucleases. Cell *91*, 457–466.

Mitchell, P., Petfalski, E., and Tollervey, D. (1996). The 3′ end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. Genes Dev. *10*, 502–513.

Mitchell, S. F., Jain, S., She, M., and Parker, R. (2013). Global analysis of yeast mRNPs. Nat. Struct. Mol. Biol. *20*, 127–133.

Mitchell, S. F., and Parker, R. (2014). Principles and properties of eukaryotic mRNPs. Mol. Cell *54*, 547–558.

Montpetit, B., Thomsen, N. D., Helmke, K. J., Seeliger, M. A., Berger, J. M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature *472*, 238–244.

Moorefield, B., Greene, E. A., and Reeder, R. H. (2000). RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. Proc. Natl. Acad. Sci. U. S. A. *97*, 4724–4729.

Moreau, K., Le Dantec, A., Mosrin-Huaman, C., Bigot, Y., Piégu, B., and Rahmouni, A. R.

(2019). Perturbation of mRNP biogenesis reveals a dynamic landscape of the Rrp6-dependent surveillance machinery trafficking along the yeast genome. RNA Biol. *16*, 879–889.
Morton, D. J., Kuiper, E. G., Jones, S. K., Leung, S. W., Corbett, A. H., and Fasken, M. B. (2018). The RNA exosome and RNA exosome-linked disease. RNA *24*, 127–142.
Moteki, S., and Price, D. (2002). Functional coupling of capping and transcription of mRNA. Mol. Cell *10*, 599–609.
Müller-McNicoll, M., and Neugebauer, K. M. (2013). How cells get the message: dynamic assembly and function of mRNA-protein complexes. Nat. Rev. Genet. *14*, 275–287.
Murphy, R., and Wente, S. R. (1996). An RNA-export mediator with an essential nuclear export signal. Nature *383*, 357–360.
Nasmyth, K., and Haering, C. H. (2005). The structure and function of SMC and kleisin complexes. Annu. Rev. Biochem. *74*, 595–648.
Neil, H., Malabat, C., D'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L. M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038–1042.
Neumann, S., Petfalski, E., Brügger, B., Grosshans, H., Wieland, F., Tollervey, D., and Hurt, E. (2003). Formation and nuclear export of tRNA, rRNA and mRNA is regulated by the ubiquitin ligase Rsp5p. EMBO Rep. *4*, 1156–1162.
Ng, R., and Abelson, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *77*, 3912–3916.
Niño, C. A., Hérissant, L., Babour, A., and Dargemont, C. (2013). mRNA nuclear export in yeast. Chem. Rev. *113*, 8523–8545.
Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920–930.
Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J. Biol. Chem. *282*, 15645–15651.
van Nues, R., Schweikert, G., de Leau, E., Selega, A., Langford, A., Franklin, R., Iosub, I., Wadsworth, P., Sanguinetti, G., and Granneman, S. (2017). Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. Nat. Commun. *8*, 12.
Oeffinger, M., and Montpetit, B. (2015). Emerging properties of nuclear RNP biogenesis and

export. Curr. Opin. Cell Biol. *34*, 46–53.

Oeffinger, M., Wei, K. E., Rogers, R., DeGrasse, J. A., Chait, B. T., Aitchison, J. D., and Rout, M. P. (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. Nat. Methods *4*, 951–956.

Oeffinger, M., and Zenklusen, D. (2012). To the pore and through the pore: a story of mRNA export kinetics. Biochim. Biophys. Acta *1819*, 494–506.

Pak, C. H. *et al.* (2011). Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in Drosophila and humans. Proc. Natl. Acad. Sci. U. S. A. *108*, 12390–12395.

Palancade, B., Zuccolo, M., Loeillet, S., Nicolas, A., and Doye, V. (2005). Pml39, a novel protein of the nuclear periphery required for nuclear retention of improper messenger ribonucleoparticles. Mol. Biol. Cell *16*, 5258–5268.

Palmer, D. K., O'Day, K., Trong, H. L., Charbonneau, H., and Margolis, R. L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc. Natl. Acad. Sci. U. S. A. *88*, 3734–3738.

Parker, S. *et al.* (2017). A resource for functional profiling of noncoding RNA in the yeast Saccharomyces cerevisiae. RNA *23*, 1166–1171.

Parker, S. *et al.* (2018). Large-scale profiling of noncoding RNA function in yeast. PLOS Genet. *14*, e1007253.

Passmore, L. A., Schmeing, T. M., Maag, D., Applefield, D. J., Acker, M. G., Algire, M. A. A., Lorsch, J. R., and Ramakrishnan, V. (2007). The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome. Mol. Cell *26*, 41–50.

Paul, B., and Montpetit, B. (2016). Altered RNA processing and export lead to retention of mRNAs near transcription sites and nuclear pore complexes or within the nucleolus. Mol. Biol. Cell *27*, 2742–2756.

Peña, Á. *et al.* (2012). Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. EMBO J. *31*, 1605–1616.

Piekna-Przybylska, D., Decatur, W. A., and Fournier, M. J. (2007). New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. RNA *13*, 305–312.

Porrua, O., Boudvillain, M., and Libri, D. (2016). Transcription Termination: Variations on Common Themes. Trends Genet. *32*, 508–522.

Porrua, O., and Libri, D. (2013). RNA quality control in the nucleus: The Angels' share of RNA. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 604–611.
Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nat. Rev. Mol. Cell Biol. *16*, 190–202.
Preker, P. J., Ohnacker, M., Minvielle-Sebastia, L., and Keller, W. (1997). A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. EMBO J. *16*, 4727–4737.
Del Priore, V., Snay, C. A., Bahr, A., and Cole, C. N. (1996). The product of the Saccharomyces cerevisiae RSS1 gene, identified as a high-copy suppressor of the rat7-1 temperature-sensitive allele of the RAT7/NUP159 nucleoporin, is required for efficient mRNA export. Mol. Biol. Cell *7*, 1601–1621.
Proudfoot, N. (2004). New perspectives on connecting messenger RNA 3′ end formation to transcription. Curr. Opin. Cell Biol. *16*, 272–278.
Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501–512.
Reeder, R. H., Guevara, P., and Roan, J. G. (1999). Saccharomyces cerevisiae RNA polymerase I terminates transcription at the Reb1 terminator in vivo. Mol. Cell. Biol. *19*, 7369–7376.
Reis, C. C., and Campbell, J. L. (2007). Contribution of Trf4/5 and the nuclear exosome to genome stability through regulation of histone mRNA levels in Saccharomyces cerevisiae. Genetics *175*, 993–1010.
Rialdi, A. *et al.* (2017). The RNA Exosome Syncs IAV-RNAPII Transcription to Promote Viral Ribogenesis and Infectivity. Cell *169*, 679-692.e14.
Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
Rodriguez, C. R., Takagi, T., Cho, E. J., and Buratowski, S. (1999). A saccharomyces cerevisiae RNA 5'-triphosphatase related to mRNA capping enzyme. Nucleic Acids Res. *27*, 2181–2188.
Rodriguez, M. S., Gwizdek, C., Haguenauer-Tsapis, R., and Dargemont, C. (2003). The HECT ubiquitin ligase Rsp5p is required for proper nuclear export of mRNA in Saccharomyces cerevisiae. Traffic *4*, 566–575.
Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224*, 234–237.

Roeder, R. G., and Rutter, W. J. (1970). Specific nucleolar and nucleoplasmic RNA polymerases. Proc. Natl. Acad. Sci. U. S. A. *65*, 675–682.

Rohner, S., Gasser, S. M., and Meister, P. (2008). Modules for cloning-free chromatin tagging in Saccharomyces cerevisae. Yeast *25*, 235–239.

Rondón, A. G., Jimeno, S., and Aguilera, A. (2010). The interface between transcription and mRNP export: From THO to THSC/TREX-2. Biochim. Biophys. Acta - Gene Regul. Mech. *1799*, 533–538.

Rougemaille, M. *et al.* (2008). THO/Sub2p functions to coordinate 3'-end processing with gene-nuclear pore association. Cell *135*, 308–321.

Rougemaille, M., Gudipati, R. K., Olesen, J. R., Thomsen, R., Seraphin, B., Libri, D., and Jensen, T. H. (2007). Dissecting mechanisms of nuclear mRNA surveillance in THO/sub2 complex mutants. EMBO J. *26*, 2317–2326.

Ryan, K., Calvo, O., and Manley, J. L. (2004). Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. RNA *10*, 565–573.

Saavedra, C., Tung, K. S., Amberg, D. C., Hopper, A. K., and Cole, C. N. (1996). Regulation of mRNA export in response to stress in Saccharomyces cerevisiae. Genes Dev *10*, 1608–1620.

Sadowski, M., Dichtl, B., Hübner, W., and Keller, W. (2003). Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. EMBO J. *22*, 2167–2177.

Saguez, C., Schmid, M., Olesen, J. R., Ghazy, M. A. E.-H., Qu, X., Poulsen, M. B., Nasser, T., Moore, C., and Jensen, T. H. (2008). Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. Mol. Cell *31*, 91–103.

Saitoh, Y.-H., Ogawa, K., and Nishimoto, T. (2005). Brl1p -- a novel nuclear envelope protein required for nuclear transport. Traffic *6*, 502–517.

Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., and Burge, C. B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science *320*, 1643–1647.

Santaguida, S., and Amon, A. (2015). Short- and long-term effects of chromosome mis-segregation and aneuploidy. Nat. Rev. Mol. Cell Biol. *16*, 473–485.

Santos-Rosa, H., Moreno, H., Simos, G., Segref, A., Fahrenkrog, B., Pante, N., and Hurt, E. (1998). Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the

nuclear pores. Mol Cell Biol *18*, 6826–6838.

Saroufim, M.-A., Bensidoun, P., Raymond, P., Rahman, S., Krause, M. R., Oeffinger, M., and Zenklusen, D. (2015). The nuclear basket mediates perinuclear mRNA scanning in budding yeast. J. Cell Biol. *211*, 1131–1140.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F. P., Dastidar, E. G., Sanchez-Rotunno, M., Arraiano, C. M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat. Struct. Mol. Biol. *16*, 56–62.

Schilders, G., Raijmakers, R., Raats, J. M. H., and Pruijn, G. J. M. (2005). MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. Nucleic Acids Res. *33*, 6795–6804.

Schindelin, J. *et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2003). Structures of deacylated tRNA mimics bound to the E site of the large ribosomal subunit. RNA *9*, 1345–1352.

Schmid, M., Olszewski, P., Pelechano, V., Gupta, I., Steinmetz, L. M., and Jensen, T. H. (2015). The Nuclear PolyA-Binding Protein Nab2p Is Essential for mRNA Production. Cell Rep. *12*, 128–139.

Schmid, M., Poulsen, M. B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L. M., Moore, C., and Jensen, T. H. (2012). Rrp6p Controls mRNA Poly(A) Tail Length and Its Decoration with Poly(A) Binding Proteins. Mol. Cell *47*, 267–280.

Schmidt, C. *et al.* (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science *354*, 1431–1433.

Schmidt, K., and Butler, J. S. (2013). Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. Wiley Interdiscip. Rev. RNA *4*, 217–231.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide analysis of exosome targets. Mol. Cell *48*, 422–433.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res. *37*, 1127–1140.

Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. Trends

Biochem. Sci. *38*, 485–493.
Schneider, D. A. (2012). RNA polymerase I activity is regulated at multiple steps in the transcription cycle: Recent insights into factors that influence transcription elongation. Gene *493*, 176–184.
Schneider, R., and Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. Genes Dev. *21*, 3027–3043.
Schuch, B., Feigenbutz, M., Makino, D. L., Falk, S., Basquin, C., Mitchell, P., and Conti, E. (2014). The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. EMBO J. *33*, 2829–2846.
Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). XTranscriptome surveillance by selective termination of noncoding RNA synthesis. Cell *155*, 1075–1087.
Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A., and Mühlemann, O. (2013). Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim. Biophys. Acta *1829*, 612–623.
Scott, D. D. *et al.* (2017). Nol12 is a multifunctional RNA binding protein at the nexus of RNA and DNA metabolism. Nucleic Acids Res. *45*, 12509–12528.
Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Lührmann, R., and Hurt, E. (1997). Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. EMBO J. *16*, 3256–3271.
Selmer, M., Dunham, C. M., Murphy, F. V., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. Science *313*, 1935–1942.
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.
Sheff, M. A., and Thorn, K. S. (2004). Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. Yeast *21*, 661–670.
Sheltzer, J. M., Torres, E. M., Dunham, M. J., and Amon, A. (2012). Transcriptional consequences of aneuploidy. Proc. Natl. Acad. Sci. U. S. A. *109*, 12644–12649.
Shibagaki, Y., Itoh, N., Yamada, H., Nagata, S., and Mizumoto, K. (1992). mRNA capping

enzyme: Isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from Saccharomyces cerevisiae. J. Biol. Chem. *267*, 9521–9528.

Shiokawa, K., and Pogo, A. O. (1974). The role of cytoplasmic membranes in controlling the transport of nuclear messenger RNA and initiation of protein synthesis. Proc. Natl. Acad. Sci. U. S. A. *71*, 2658–2662.

Sikorski, R. S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Singh, G., Pratt, G., Yeo, G. W., and Moore, M. J. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. Annu. Rev. Biochem. *84*, 325–354.

Siomi, H., Matunis, M. J., Michael, W. M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionary conserved motif. Nucleic Acids Res. *21*, 1193–1198.

Smith, C. *et al.* (2015). In vivo single-particle imaging of nuclear mRNA export in budding yeast demonstrates an essential role for Mex67p. J. Cell Biol. *211*, 1121–1130.

Snay-Hodge, C. a, Colot, H. V, Goldstein, a L., and Cole, C. N. (1998). Dbp5p/Rat8p is a yeast nuclear pore-associated DEAD-box protein essential for RNA export. EMBO J. *17*, 2663–2676.

Soheilypour, M., and Mofrad, M. R. K. (2016). Regulation of RNA-binding proteins affinity to export receptors enables the nuclear basket proteins to distinguish and retain aberrant mRNAs. Sci. Rep. *6*, 35380.

Sohrabi-Jahromi, S., Hofmann, K. B., Boltendahl, A., Roth, C., Gressel, S., Baejen, C., Soeding, J., and Cramer, P. (2019). Transcriptome maps of general eukaryotic RNA degradation factors. Elife *8*.

Sosnowski, P., Urnavicius, L., Boland, A., Fagiewicz, R., Busselez, J., Papai, G., and Schmidt, H. (2018). The CryoEM structure of the Saccharomyces cerevisiae ribosome maturation factor Rea1. Elife *7*.

Spahn, C. M. T., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G., and Frank, J. (2001). Structure of the 80S ribosome from Saccharomyces cerevisiae - tRNA-ribosome and subunit-subunit interactions. Cell *107*, 373–386.

Spies, N., Burge, C. B., and Bartel, D. P. (2013). 3′ UTR-Isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Res. *23*, 2078–2090.

Staub, E., Mackowiak, S., and Vingron, M. (2006). An inventory of yeast proteins associated

with nucleolar and ribosomal components. Genome Biol. *7*, R98.
Stead, J. A., Costello, J. L., Livingstone, M. J., and Mitchell, P. (2007). The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. Nucleic Acids Res. *35*, 5556–5567.
Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature *413*, 327–331.
Strabetaer, K., Sträßer, K., Hurt, E., Strässer K, and Hurt E (2000). Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. EMBO J. *19*, 410–420.
Straight, A. F., Belmont, A. S., Robinett, C. C., and Murray, A. W. (1996). GFP-tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr. Biol. *6*, 1599–1608.
Strasser, K. *et al.* (2002). TREX is a conserved complex coupling transcription with messenger RNA export. Nature *417*, 304-8.
Sträßer, K., Baßler, J., Hurt, E., Strasser, K., and Bassler, J. (2000). Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. J Cell Biol *150*, 695–706.
Strässer, K., Hurt, E., and Strasser, K. (2001). Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature *413*, 648-52.
Stuparevic, I., Mosrin-Huaman, C., Hervouet-Coste, N., Remenaric, M., and Rahmouni, A. R. (2013). 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 (mRN. J. Biol. Chem. *288*, 31816–31829.
Subtelny, A. O., Eichhorn, S. W., Chen, G. R., Sive, H., and Bartel, D. P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66–71.
Sullivan, M., Higuchi, T., Katis, V. L., and Uhlmann, F. (2004). Cdc14 Phosphatase Induces rDNA Condensation and Resolves Cohesin-Independent Cohesion during Budding Yeast Anaphase. Cell *117*, 471–482.
Sun, M., Schwalb, B., Pirkl, N., Maier, K. C., Schenk, A., Failmezger, H., Tresch, A., and

Cramer, P. (2013). Global analysis of Eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. Mol. Cell *52*, 52–62.

Tachibana, T., Imamoto, N., Seino, H., Nishimoto, T., and Yoneda, Y. (1994). Loss of RCC1 leads to suppression of nuclear protein import in living cells. J Biol Chem *269*, 24542–24545.

Taylor, L. L., Jackson, R. N., Rexhepaj, M., King, A. K., Lott, L. K., van Hoof, A., and Johnson, S. J. (2014). The Mtr4 ratchet helix and arch domain both function to promote RNA unwinding. Nucleic Acids Res. *42*, 13861–13872.

Terry, L. J., and Wente, S. R. (2007). Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. J. Cell Biol. *178*, 1121–1132.

Thebault, P., Boutin, G., Bhat, W., Rufiange, A., Martens, J., and Nourani, A. (2011). Transcription Regulation by the Noncoding RNA SRG1 Requires Spt2-Dependent Chromatin Deposition in the Wake of RNA Polymerase II. Mol. Cell. Biol. *31*, 1288–1300.

Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M. W., and Kulozik, A. E. (1998). Binary specification of nonsense codons by splicing and cytoplasmic translation. EMBO J. *17*, 3484–3494.

Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Mol. Cell *23*, 853–864.

Thiry, M., and Lafontaine, D. L. J. (2005). Birth of a nucleolus: The evolution of nucleolar compartments. Trends Cell Biol. *15*, 194–199.

Thoms, M., Thomson, E., Baßler, J., Gnädig, M., Griesel, S., and Hurt, E. (2015). The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. Cell *162*, 1029–1038.

Thomsen, R. (2003). Localization of nuclear retained mRNAs in Saccharomyces cerevisiae. RNA *9*, 1049–1057.

Thomsen, R., Saguez, C., Nasser, T., and Jensen, T. H. (2008). General, rapid, and transcription-dependent fragmentation of nucleolar antigens in S. cerevisiae mRNA export mutants. RNA *14*, 706–716.

Torres, E. M., Sokolsky, T., Tucker, C. M., Chan, L. Y., Boselli, M., Dunham, M. J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science *317*, 916–924.

Tran, E. J., Zhang, X., and Maxwell, E. S. (2003). Efficient RNA 2'-O-methylation requires

juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. EMBO J. *22*, 3930–3940.

Tran, E. J., Zhou, Y., Corbett, A. H., and Wente, S. R. (2007). The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. Mol. Cell *28*, 850–859.

Tsukamoto, T., Shibagaki, Y., Imajoh-Ohmi, S., Murakoshi, T., Suzuki, M., Nakamura, A., Gotoh, H., and Mizumoto, K. (1997). Isolation and characterization of the yeast mRNA capping enzyme β subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. Biochem. Biophys. Res. Commun. *239*, 116–122.

Tuck, A. C., and Tollervey, D. (2013). A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. Cell *154*, 996–1009.

Tudek, A. *et al.* (2014). Molecular Basis for Coordinating Transcription Termination with Noncoding RNA Degradation. Mol. Cell *55*, 467–481.

Tudek, A., Lloret-Llinares, M., and Jensen, T. H. (2018a). The multitasking polyA tail: nuclear RNA maturation, degradation and export. Philos. Trans. R. Soc. B Biol. Sci. *373*, 20180169.

Tudek, A., Schmid, M., and Jensen, T. H. (2019). Escaping nuclear decay: the significance of mRNA export for gene expression. Curr. Genet. *65*, 473–476.

Tudek, A., Schmid, M., Makaras, M., Barrass, J. D., Beggs, J. D., and Jensen, T. H. (2018b). A Nuclear Export Block Triggers the Decay of Newly Synthesized Polyadenylated RNA. Cell Rep. *24*, 2457-2467.e7.

Tycowski, K. T., Smith, C. M., Shu, M. Di, and Steitz, J. A. (1996). A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus. Proc. Natl. Acad. Sci. U. S. A. *93*, 14480–14485.

Vanácová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. *3*, e189.

Vasiljeva, L., and Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Mol. Cell *21*, 239–248.

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Hieter, P., Vogelstein, B., and Kinzler, K. W. (1997). Characterization of the yeast transcriptome. Cell *88*, 243–251.

Veldman, G. M., Brand, R. C., Klootwijk, J., and Planta, R. (1980). Some characteristics of processing sites in ribosomal precursor RNA of yeast. Nucleic Acids Res. *8*, 2907–2920.
Vera, J. M., and Dowell, R. D. (2016). Survey of cryptic unstable transcripts in yeast. BMC Genomics *17*, 305.
Vinciguerra, P., Iglesias, N., Camblong, J., Zenklusen, D., and Stutz, F. (2005). Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. EMBO J. *24*, 813–823.
Viphakone, N., Voisinet-Hakil, F., and Minvielle-Sebastia, L. (2008). Molecular dissection of mRNA poly(A) tail length control in yeast. Nucleic Acids Res. *36*, 2418–2433.
Volkening, K., and J. Strong, M. (2011). RNA Metabolism in Neurodegenerative Disease. Curr. Chem. Biol. *5*, 90–98.
Wang, D., Mansisidor, A., Prabhakar, G., and Hochwagen, A. (2016). Condensin and Hmo1 Mediate a Starvation-Induced Transcriptional Position Effect within the Ribosomal DNA Array. Cell Rep. *14*, 1010–1017.
Wang, H.-W., Wang, J., Ding, F., Callahan, K., Bratkowski, M. A., Butler, J. S., Nogales, E., and Ke, A. (2007). Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. Proc. Natl. Acad. Sci. U. S. A. *104*, 16844–16849.
Wasmuth, E. V., and Lima, C. D. (2012). Exo- and Endoribonucleolytic Activities of Yeast Cytoplasmic and Nuclear RNA Exosomes Are Dependent on the Noncatalytic Core and Central Channel. Mol. Cell *48*, 133–144.
Wasmuth, E. V, Januszyk, K., and Lima, C. D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature *511*, 435–439.
Wassarman, D. A., and Steitz, J. A. (1992). Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. Science *257*, 1918–1925.
Watkins, N. J., and Bohnsack, M. T. (2012). The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip. Rev. RNA *3*, 397–414.
Webb, S., Hector, R. D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, R8.
Weick, E. M., Puno, M. R., Januszyk, K., Zinder, J. C., DiMattia, M. A., and Lima, C. D. (2018).

Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex. Cell *173*, 1663-1677.e21.
Weill, L., Belloc, E., Bava, F. A., and Méndez, R. (2012). Translational control by changes in poly(A) tail length: Recycling mRNAs. Nat. Struct. Mol. Biol. *19*, 577–585.
Weinmann, R., and Roeder, R. G. (1974). Role of DNA dependent RNA polymerase III in the transcription of the tRNA and 5S RNA genes. Proc. Natl. Acad. Sci. U. S. A. *71*, 1790–1794.
Weir, J. R., Bonneau, F., Hentschel, J., and Conti, E. (2010). 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. *107*, 12139–12144.
Weirich, C. S., Erzberger, J. P., Berger, J. M., and Weis, K. (2004). The N-terminal domain of Nup159 forms a beta-propeller that functions in mRNA export by tethering the helicase Dbp5 to the nuclear pore. Mol. Cell *16*, 749–760.
Weirich, C. S., Erzberger, J. P., Flick, J. S., Berger, J. M., Thorner, J., and Weis, K. (2006). Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. Nat. Cell Biol. *8*, 668–676.
Wells, S. E., Hillner, P. E., Vale, R. D., and Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Mol. Cell *2*, 135–140.
van Werven, F. J., Neuert, G., Hendrick, N., Lardenois, A., Buratowski, S., van Oudenaarden, A., Primig, M., and Amon, A. (2012). Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. Cell *150*, 1170–1181.
Wery, M., Ruidant, S., Schillewaert, S., Lepore, N., Lafontaine, D. L. J. J., Leporé, N., and Lafontaine, D. L. J. J. (2009). The nuclear poly(A) polymerase and Exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance. RNA *15*, 406–419.
Whitehouse, I., Rando, O. J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031–1035.
Woolford, J. L., and Baserga, S. J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics *195*, 643–681.
Wu, H., Yin, Q. F., Luo, Z., Yao, R. W., Zheng, C. C., Zhang, J., Xiang, J. F., Yang, L., and Chen, L. L. (2016). Unusual Processing Generates SPA LncRNAs that Sequester Multiple RNA Binding Proteins. Mol. Cell *64*, 534–548.
Wu, J., Bao, A., Chatterjee, K., Wan, Y., and Hopper, A. K. (2015). Genome-wide screen

uncovers novel pathways for tRNA processing and nuclear-cytoplasmic dynamics. Genes Dev. *29*, 2633–2644.

Wyers, F. *et al.* (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725–737.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L. M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D., and Shatkin, A. J. (1997). Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *94*, 12898–12903.

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 A resolution. Science *292*, 883–896.

Zanchin, N. I. T., and Goldfarb, D. S. (1999). The exosome subunit Rrp43p is required for the efficient maturation of 5.8S, 18S and 25S rRNA. Nucleic Acids Res. *27*, 1283–1288.

Zander, G., Hackmann, A., Bender, L., Becker, D., Lingner, T., Salinas, G., and Krebber, H. (2016). mRNA quality control is bypassed for immediate export of stress-responsive transcripts. Nature *540*, 593–596.

Zenklusen, D., Vinciguerra, P., Wyss, J. C., and Stutz, F. (2002). Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. Mol Cell Biol *22*, 8241-53.

Zhao, J., Kessler, M., Helmling, S., O'Connor, J. P., and Moore, C. (1999). Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. Mol. Cell. Biol. *19*, 7733–7740.

Zinder, J. C. C., Wasmuth, E. V. V, and Lima, C. D. D. (2016). Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3. Mol. Cell *64*, 1–12.

Zinder, J. C., and Lima, C. D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev. *31*, 88–100.

Appendices

Appendix A

Table 6-1: List of mutant alleles with a reported poly(A)-RNA accumulation phenotype within the temperature sensitive mutant collection

Gene	Identified in screen	Description	Reference
brl1-3231	Yes	Nuclear envelope membrane protein; functions with Apq12p and Brr6p in lipid homeostasis	(Saitoh *et al.*, 2005)
brr6-1	Yes	Nuclear envelope membrane protein; functions with Apq12p and Brl1p in lipid homeostasis	(de Bruyn Kops and Guthrie, 2001)
dbp5-1	Yes	DEAD-box ATPase; involved in mRNA export from the nucleus	(Snay-Hodge *et al.*, 1998)
dis3-1	Yes	Exosome core complex catalytic subunit; involved in 3'-5' RNA processing and degradation	(Kadowaki *et al.*, 1994a)
gle1-4	Yes	Cytoplasmic nucleoporin; required for polyadenylated mRNA export	(Murphy and Wente, 1996)
ipl1-1	Yes	Aurora kinase of chromosomal passenger complex; mediates release of mono-oriented kinetochores	(Cole *et al.*, 2002)
mex67-5	Yes	RNA binding protein; involved in nuclear mRNA export	(Segref *et al.*, 1997)
mtr4-1	Yes	3'-5' RNA helicase; involved in nuclear RNA processing and decay as part of the TRAMP complex	(Kadowaki *et al.*, 1994a)
nup145-4	Yes	Structural component of the NPC; involved in NPC biogenesis and genes tethering to the nuclear periphery	(Fabre *et al.*, 1994)
nup159-1	Yes	FG-nucleoporin that is part of NPC cytoplasmic filaments; Dbp5 regulator in mRNA export	(Del Priore *et al.*, 1996)
pcf11-1	No	mRNA 3' end processing factor; essential component of cleavage and polyadenylation factor IA	(Hammell *et al.*, 2002)
pta1-1	Yes	Cleavage and polyadenylation factor complex subunit; required for processing mRNA and snoRNA 3' ends	(Hammell *et al.*, 2002)
rat1-1	No	5' to 3' RNA exonuclease; involved in ncRNA processing and mRNA transcription termination	(Amberg *et al.*, 1992)
rat9-1	No	Structural subunit of NPCs; contributes to nucleocytoplasmic transport and NPC biogenesis	(Goldstein *et al.*, 1996)
rna15-58	No	Cleavage and polyadenylation factor I subunit; involved in processing mRNA 3' ends	(Hammell *et al.*, 2002)
rsp5-3	Yes	E3 ubiquitin ligase; regulates MVB sorting, the heat shock response, transcription, and ribosome stability	(Neumann *et al.*, 2003)
srm1-ts	Yes	Nucleotide exchange factor for Gsp1p; required for nucleocytoplasmic trafficking through NPCs	(Kadowaki *et al.*, 1994a)

Appendix B

R codes for RNA-seq analyses are given below. You might need to generate a bam file to start this analysis.

R code for running featureCount function on all bam files generated from HISAT2

```
setwd("/home/biplab/DATA1/fastq_files/")
myFiles <- list.files(pattern="*bam")
library("Rsubread")
gene_csl4 <- featureCounts(files<-myFiles,

            # annotation
       annot.inbuilt=NULL,
       annot.ext= ~/project_exosome/input_files/gtf_files/
       Saccharomyces_cerevisiae.R64-1-1.85.gtf,
       isGTFAnnotationFile=TRUE,
       GTF.featureType= "gene",
                      GTF.attrType="gene_id",
                         chrAliases=NULL,

                         # level of summarization
                         useMetaFeatures=FALSE,

                         # overlap between reads and features
                         allowMultiOverlap=TRUE,
                         minOverlap=10,
                         largestOverlap=FALSE,
                         readExtension5=0,
                         readExtension3=0,
                         read2pos=NULL,

                         # multi-mapping reads
                         countMultiMappingReads=TRUE,
                         fraction=TRUE,

                         # read filtering
                         minMQS=0,
                         splitOnly=FALSE,
                         nonSplitOnly=FALSE,
                         primaryOnly=FALSE,
                         ignoreDup=FALSE,

                         # strandness
                         strandSpecific=2,

                         # exon-exon junctions
```

```
                          juncCounts=TRUE,
                          genome=NULL,

                        # parameters specific to paired end reads
                          isPairedEnd=TRUE,
                          requireBothEndsMapped=TRUE,
                          checkFragLength=TRUE,
                          minFragLength=20,
                          maxFragLength=200,
                          countChimericFragments=FALSE,
                          autosort=TRUE,

                          # miscellaneous
                          nthreads=1,
                          maxMOp=10,
                          reportReads=FALSE)

gene_csl4_enp1 <- data.frame(gene_csl4_enp1$annotation,
gene_csl4_enp1$counts,stringsAsFactors=FALSE)
write.table(cuts_xu, file = "~/project_exosome/input_files/
 expression_data/fc_genes_csl4_enp1_srm1.csv", sep = "\t",
 quote = FALSE, col.names = TRUE)
```

R code for running DESeq2

```
options( warn = -1 )
library('DESeq2')
library('data.table')
library('stats')
library('ggplot2')
library('tidyverse')
setwd("~/Desktop/Paul_et_al_2019/")
wt_vs_mutants_d <- fread("./results/output_feature_count.csv",
header = T, sep = ",")
#---------------------------------- DESEQ2 RB -------------------
------------------------------

#Choosing columns for RB data
sample_name = c('GeneID', 'Rrp6_1', 'Rrp6_2', 'Dis3_1',
'Dis3_2', 'Csl4_90_1_rb', 'Csl4_90_2_rb', 'Csl4_90_3_rb',
                'Enp1_90_1_rb', 'Enp1_90_2_rb','Srm1_90_1_rb',
                'Srm1_90_2_rb', 'wt_90_1', 'wt_90_2' )
wt_vs_mutants_d <-
wt_vs_mutants_d[!duplicated(wt_vs_mutants_d$GeneID),]
colnames(wt_vs_mutants_d) <- sample_name

sample_name = c('Rrp6_1', 'Rrp6_2', 'Dis3_1', 'Dis3_2',
'Csl4_90_1_rb', 'Csl4_90_2_rb', 'Csl4_90_3_rb',
```

```
                'Enp1_90_1_rb', 'Enp1_90_2_rb','Srm1_90_1_rb',
                 'Srm1_90_2_rb', 'wt_90_1', 'wt_90_2' )


#Selecting data
expr_rb = wt_vs_mutants_d[, sample_name, with=FALSE]


#Setting up colData
col_rb = data.frame(sample = sample_name,
            treatment = c('rrp6_90','rrp6_90',
             'dis3_90', 'dis3_90', 'csl4_90',
             'csl4_90','csl4_90', 'enp1_90', 'enp1_90',
                 'srm1_90', 'srm1_90','Ctrl', 'Ctrl'))

col_rb$treatment <- factor(col_rb$treatment,
levels=c('rrp6_90', 'dis3_90', 'csl4_90',
'enp1_90', 'srm1_90', 'Ctrl'))



dds_rb <- DESeqDataSetFromMatrix(countData = data.frame(expr_rb),
                       colData = col_rb,
                       design = ~ treatment)
rownames(dds_rb) <- wt_vs_mutants_d$GeneID

#Applying the model
dds_rb <- estimateSizeFactors(dds_rb)
dds_rb <- estimateDispersions(dds_rb)
#res_lrt <- DESeq(dds_rb, test='LRT', reduced = ~1)
res_wald <- DESeq(dds_rb, test='Wald')

#getting results ...
res_rrp6_90 <- results(res_wald, pAdjustMethod = "BH",
contrast=c('treatment', 'rrp6_90', 'Ctrl'))
res_dis3_90 <- results(res_wald, pAdjustMethod = "BH",
contrast=c('treatment', 'dis3_90', 'Ctrl'))
res_csl4_90 <- results(res_wald, pAdjustMethod = "BH",
contrast=c('treatment', 'csl4_90', 'Ctrl'))
res_enp1_90 <- results(res_wald, pAdjustMethod = "BH",
contrast=c('treatment', 'enp1_90', 'Ctrl'))
res_srm1_90 <- results(res_wald, pAdjustMethod = "BH",
contrast=c('treatment', 'srm1_90', 'Ctrl'))

#saving results for RB data
tmp <- cbind(as.data.frame(res_rrp6_90),
```

```
            as.data.frame(res_dis3_90),
            as.data.frame(res_csl4_90),
            as.data.frame(res_enp1_90),
            as.data.frame(res_srm1_90))

log2fc <- tmp[ , grepl( "log2FoldChange" , names( tmp) ) ]
p_v <- tmp[ , grepl( "pvalue" , names( tmp ) ) ]
p_v <- p_v[apply(p_v[, -1], MARGIN = 1,
function(x, any(x < 0.01)), ]
new_log2fc <- log2fc[rownames(p_v),]
colnames(new_log2fc) = c('rrp6', 'dis3', 'csl4', 'enp1', 'srm1')
new_log2fc<- new_log2fc %>% rownames_to_column("gene_id")
write.table(new_log2fc, file="./results/log2FC_all_rb.tsv",
 quote = F, sep='\t', row.names = TRUE, col.names = TRUE)
```

R Code for Figure 4-1B

```
library(ggplot2)
library(RColorBrewer)
library(gplots)
library(DESeq2)
sampleinfo <- read.csv("~/Desktop/sample_info.csv", sep = ",")
df <- read.csv("~/Paul_et_al_2019/results/
output_feature_count.csv", sep = ",")
df <- df[-1]
colnames(df) <- c("rrp6_R1", "rrp6_R2", "dis3_R1", "dis3_R2",

 "csl4ph_R1", "csl4ph_R2", "csl4ph_R3", "enp1_R1",

                  "enp1_R2", "srm1_R1", "srm1_R2", "WT_R1",
"WT_R2")

ddsFullCountTable <- DESeqDataSetFromMatrix(countData = df,
colData = sampleinfo, design = ~sample)
dds <- DESeq(ddsFullCountTable)
rld <- rlog( dds )
Colors=rev(brewer.pal(11,"Spectral"))
Colors=colorRampPalette(Colors)(100)
heatmap.2(assay(rld), scale = "row", trace="none", col=Colors)
```

Figure 4-2B

```
#Code for making heatmap of correlation co-effient
library(data.table)
library(corrplot)
library(dplyr)
```

```
setwd("~/Desktop/Paul_et_al_2019/")
fc_df_rb <- read.csv("./results/log2FC_all_rb.tsv", sep='\t')
log2fc <- fc_df_rb[complete.cases(fc_df_rb), ]
df <- log2fc[, c("rrp6", "dis3", "csl4","enp1","srm1")]
M <- cor(df)
png("./Figures/corr_plot_logfc_rb.png")
corrplot(M, method = "color", type="upper", addCoef.col =
"white")
dev.off()
```

R code for Figure 4-2C

```
library(data.table)
library(ComplexHeatmap)
library(circlize)
library(ggplot2)
library(ggdendro)
library(ggpubr)
setwd("~/Desktop/Paul_et_al_2019/")

fc_df_rb <- read.csv("./results/log2FC_all_rb.tsv", sep='\t')
transcript = fread('./annotation/yeast_all_annot.gtf')

#-----------------------------------------------------------------
------------------
type_rb = c('rrp6', 'dis3', 'csl4', 'enp1', 'srm1')
data_rb = fc_df_rb[,1:6]
colnames(data_rb) <- append(c('gene_id'), type_rb)

data_rb <- as.data.frame(data_rb)
biotype = c()
slc_row = c()
for (row in 1:nrow(data_rb)){
    current_id <- data_rb[row, 'gene_id']
    to_add = transcript[transcript$gene_id ==
current_id]$gene_biotype
    gene = transcript[transcript$gene_id == current_id]$gene_id
    biotype <- c(biotype, to_add)
    slc_row <- c(slc_row, gene)
}

data_rb <- data_rb[data_rb$gene_id %in% slc_row,]


data_rb$biotype <- biotype
data_rb <- data_rb[,c(1,6,5,2,3,4,7)]
#-----------------------------------------------------------
```

```
ha = HeatmapAnnotation(df = data.frame(type = type_rb))

tmp <- data_rb[data_rb$biotype %in% c('pervasive',
'protein_coding'),]

colnames(tmp) <- c('gene_id', 'rrp6', 'dis3', 'csl4',
'enp1', 'srm1', 'RNAclass')

tmp$RNAclass[tmp$RNAclass == 'pervasive'] <- "Pervasive"
tmp$RNAclass[tmp$RNAclass == 'protein_coding'] <- "mRNA"
to_plot = as.matrix(tmp[,2:6])
rownames(to_plot) <- tmp$gene_id

kclus <- kmeans(to_plot, 6)
split <- paste0("C", kclus$cluster)

ht = Heatmap(to_plot, split=split, name = 'log2FC', gap = unit(2,
"mm"),
       col = colorRamp2(c(-5, 0, 5), c("blue", "white", "red")),
       width = unit(6,'cm'),
     row_title='Kmean clusters',
       show_row_names = FALSE, show_column_names = TRUE,
       cluster_rows = FALSE,
       cluster_columns = FALSE)

ha = Heatmap(tmp$RNAclass, name = "", show_row_names = FALSE,
        width = unit(5, "mm"), col = c('#808080','#70B2FF'))

ht_list = ht + ha
png("./Figures/figure1d.png")
draw(ht_list, row_title='Kmean clusters')
dev.off()
```

Figure 4-2D

```
library(ggplot2)
setwd("~/Desktop/Paul_et_al_2019/")
fc_df_rb <- read.csv("./results/log2FC_all_rb.tsv", sep='\t')
rna_class <- read.csv("./annotation/rna3.csv", sep = "\t")
xut <- read.csv("./annotation/xut_list.txt", header = F, sep =
";")
colnames(xut) <- c("gene_id", "RNAclass")
rna_class <- rbind(rna_class, xut)
df <- merge(fc_df_rb, rna_class, by = 'gene_id', all=FALSE)
df <- df[apply(df!=0, 1, all),]
```

```
df <- df[,-1]
df1 <- melt(df)
png("./Figures/figure1e.png")
ggplot(df1, aes(x=value, color=RNAclass)) +
geom_line(stat="density",  size = 1.0) +
facet_wrap(~variable) +
xlim(-4, 8) + theme_bw()
dev.off()
```

Figure 4-4A

```
library(ggplot2)
library(tidyverse)
fc_df_rb <- read.csv("./results/log2FC_all_rb.tsv", sep='\t')
rna_class <- read.csv("./annotation/tollarvy_rna_classes2.csv",
sep = "\t")
colnames(rna_class) <- c("gene_id", "RNAclass")
df <- merge(fc_df_rb, rna_class, by = 'gene_id', all=FALSE)

df <- df[,-1]
df1 <- melt(df)
png("./Figures/figure3a.png")
ggplot(df1, aes(x = RNAclass, y = value, fill=variable)) +
 facet_wrap( ~ variable) + xlab("Classes of mRNA") +
  ylab("log2FC") + ylim(-6,6) + geom_boxplot(outlier.size
  =0.1,notch=TRUE) +
   theme(panel.background = element_rect(fill = "white"),
  axis.line = element_line(size = 0.7, color = "black"),
  axis.text=element_text(size=8),
axis.title=element_text(size=14,
  face="bold"),
strip.text.x = element_text(size = 15,
face ="bold.italic"), legend.position = c(0.8, 0.2))
dev.off()
```

Shell code for aligning all RNA-Seq read to yeast genome.

```
for f in `ls *.fastq | sed 's/_R[12].fastq//g' | sort -u`

do

hisat2 -p 4 --fr -x ~/R64-1-1/Sequence/WholeGenomeFasta/genome
--known-splicesite-infile ~/yeast_splice_sites.txt
-1 ${f}_R1.fastq -2 ${f}_R2.fastq |
samtools view -Sb > ~/DATA2/hisat2_out/${f}.bam
```