Title
Insights into RNA Splicing and the Regulation of Gene Expression in Saccharomyces cerevisiae

Author
Gabunilas, Jason

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNVERSITY OF CALIFORNIA

Los Angeles

Insights into RNA Splicing and the Regulation of Gene Expression in *Saccharomyces cerevisiae*

A dissertation submitted in partial satisfaction of

the requirements for the degree Doctor of Philosophy

in Biochemistry and Molecular Biology

by

Jason Anthony Gabunilas

2017

ABSTRACT OF THE DISSERTATION

Insights into RNA Splicing and the Regulation of Gene Expression

in *Saccharomyces cerevisiae*

by

Jason Anthony Gabunilas

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2017

Professor Guillaume Chanfreau, Chair

RNA splicing is a critical component in the regulation of gene expression in all eukaryotes. The work described herein chronicles our investigative efforts into three facets of RNA splicing and their associated mechanisms in the model organism *Saccharomyces cerevisiae*. Previous work from our group highlighted the ability for nonsense-mediated mRNA decay (NMD) to mask the splicing defects of splicing factor mutants, suggesting that the full repertoire of splicing substrates and products is at least partly occluded by RNA surveillance mechanisms. Continuing this work, we sought to uncover previously unidentified splicing events by performing RNA-sequencing in wild-type yeast as well as strains deficient in NMD. This analysis revealed that alternative splicing at unannotated non-canonical 5'- and 3'-splice sites occurs within a large number pre-mRNAs in yeast, but that these events are not usually observed because they introduce premature termination codons (PTCs) into the translational reading frames of the spliced transcripts, thereby rendering them targets for degradation by NMD. This work demonstrated that the degree of alternative splicing in yeast RNA transcripts is greater than previously appreciated, and that alternative splicing linked to NMD (AS-NMD) serves to regulate overall transcript levels. Notably,

this study uncovered a non-productive alternative 5'-splice site for the ribosomal protein gene *RPL22B* that is activated in a stress-dependent manner. We further investigated the splicing of *RPL22B* and found that its protein product Rpl22p functions in an extra-ribosomal capacity by inhibiting the splicing of its own pre-mRNA, defining an autoregulatory splicing-mediated negative feedback mechanism that fine-tunes the expression of Rpl22p with potential implications in stress response. Finally, we identified a global role for the second-step splicing factor Prp18p in the suppression of non-canonical alternative 3'-splice sites throughout the yeast transcriptome. Specifically, we found that branchpoint-proximal alternative 3'-splice sites are activated in the absence of Prp18p in a substantial fraction of intron-containing genes. These results suggest that Prp18p is responsible for maintaining the fidelity of RNA splicing. Together, these studies reveal new insights into gene regulation by highlighting the interplay between RNA splicing and quality control.

The dissertation of Jason Anthony Gabunilas is approved.

Albert J. Courey

Jeffery Floyd Miller

Guillaume Chanfreau, Committee Chair

University of California, Los Angeles

2017

TABLE OF CONTENTS **PAGE**

LIST OF FIGURES **PAGE**

LIST OF SUPPLEMENTARY FIGURES **PAGE**

LIST OF SUPPLEMENTARY TEXT DOCUMENTS **PAGE**

ACKNOWLEDGEMENTS

This dissertation is the culmination of work made possible with the support, collegiality and mentorship of multiple individuals with whom I have had the pleasure of working during my time as a graduate student. Firstly, this work is the result of the guidance and mentorship of my advisor Dr. Guillaume Chanfreau, who has provided a nurturing and intellectually stimulating environment for my training. I am enormously grateful to Guillaume for offering me the opportunity to develop and mature as a scientist in his lab as well as for his unwavering support and trust in my capability to succeed, especially when I felt discouragement and uncertainty in my own abilities.

I would like to acknowledge previous graduate students Cynthia Lee, Camille Nery, and Domi Hodko as well as former post-doc Cesar Fernandez for their guidance and tutelage, especially during my initial months in the lab. I extend my heartfelt thanks to former graduate student Tad Kawashima, whose tireless work laid the foundation for my own projects in the lab and with whom I had the pleasure to publish the work presented in Chapter 2. Likewise, I would like to give special recognition to Kevin Roy, with whom I enjoyed many hours of intellectual conversation regarding our individual and collaborative projects. Specifically, my collaborations with Kevin have proven both productive and enlightening, resulting in a publication and a manuscript in preparation as presented in Chapter 4. I have immensely cherished the friendship and support of fellow graduate student Charles Wang, whose presence in the lab contributes toward a work environment that I look forward to coming to every day.

I also owe a debt of gratitude to numerous undergraduate students and lab technicians who have contributed to research efforts in the lab as well as provided much-appreciated opportunities for mentorship. I would like to acknowledge Jocelyn Almanza, Joyce Samson, Rebekah Liu, Guochang Lyu, Doug Zhang, Duy Ngo, and Richard Zhuang for their assistance. Specifically, I would like to recognize Jocelyn for her contributions to Chapter 3 and Joyce and

Guochang for their contributions to Chapter 4 of this dissertation. I would like to thank lab technicians Abby Gillespie and Tara Herrera for their key contributions to projects that I have been involved with and for helping to make the lab a truly pleasant place to work.

Beyond the laboratory, I must also acknowledge my good friend, fellow graduate student, and roommate Oscar Campos, with whom I happily spent much time discussing the trials, tribulations, and joys of life both within and outside of graduate school, explored mountains and hiking trails throughout Southern California, and developed a bond of friendship that I will treasure for years to come. Most importantly, I would like to recognize my family for their steadfast love and support over these many years in graduate school. My parents and brother have been my greatest supporters as I have strived to achieve my goals. Their dedication to my endeavors serve as a constant reminder that regardless of where my career takes me, my true home will always be with them.

Reprint of Publications

Kawashima T, Douglass S, **Gabunilas J**, Pellegrini M, Chanfreau GF (2014) Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*. PLoS Genet 10(4): e1004249. doi:10.1371/journal.pgen.1004249

Chapter 2 is a reprint of this publication, permitted under Creative Commons Attribution 4.0 International (CC BY 4.0). My contributions to this work include the Northern blots presented in Figure 5 and assistance with manuscript revision. First author Tadashi Kawashima performed the majority of experiments and wrote the manuscript with principal investigator Guillaume Chanfreau. Co-authors Stephen Douglass and Matteo Pellegrini performed the bioinformatics analysis.

Gabunilas J, Chanfreau G (2016) Splicing-Mediated Autoregulation Modulates Rpl22p Expression in *Saccharomyces cerevisiae*. PLoS Genet 12(4): e1005999. doi:10.1371/journal.pgen.1005999

Chapter 3 is a reprint of this publication, permitted under Creative Commons Attribution 4.0 International (CC BY 4.0).

Chapter 4 is a draft of a manuscript in preparation for publication. I performed the majority of experiments and wrote the manuscript with principal investigator Guillaume Chanfreau. Joint first author Kevin Roy performed the bioinformatics analysis. Co-authors Jonelle White, Samantha Edwards, Guochang Lyu and Joyce Samson contributed to yeast strain construction, performing experiments, Northern blotting, and conducting splicing analysis by RT-PCR.

Funding Sources

- NIH Research Supplements to Promote Diversity in Health-Related Research (2012-2013)
- Ruth L. Kirschstein National Research Service Award GM007185 (2013-2016)
- UCLA Graduate Division Dissertation Year Fellowship (2016-2017)

VITA

EDUCATION

2010 B.S. in Bioengineering: Biotechnology
University of California San Diego

2011 M.S. in Bioengineering
University of California San Diego

PUBLICATIONS

Gillespie A, **Gabunilas J**, Jen JC, and Chanfreau GF. (2017) "Mutations of EXOSC3/Rrp40p Associated with Neurological Diseases Impact Ribosomal RNA Processing Functions of the Exosome in *S. cerevisiae*." RNA. doi:10.1261/rna.060004.116

Roy K, **Gabunilas J**, Gillespie A, Ngo D, and Chanfreau GF. (2016). "Common genomic elements promote transcriptional and DNA replication roadblocks." Genome Res 26(10): 1363-1375.

Gabunilas J, Chanfreau G (2016) Splicing-Mediated Autoregulation Modulates Rpl22p Expression in *Saccharomyces cerevisiae*. PLoS Genet 12(4): e1005999. doi:10.1371/journal.pgen.1005999

Kawashima T, Douglass S, **Gabunilas J**, Pellegrini M, Chanfreau GF (2014) Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*. PLoS Genet 10(4): e1004249. doi:10.1371/journal.pgen.1004249

CONFERENCE PRESENTATIONS

“Prp18 promotes global second step splicing fidelity by suppressing non-canonical 3’-splice sites.” Oral abstract presented at The 22nd Annual Meeting of the RNA Society, Prague, Czech Republic (2017 June).

“Mutations of EXOSC3/Rrp40p Associated with Neurological Diseases Impact Ribosomal RNA Processing Functions of the Exosome in *S. cerevisiae*.” Poster session presented at The 22nd Annual Meeting of the RNA Society, Prague, Czech Republic (2017 June).

“Global Suppression of non-canonical 3’-splice site usage by Prp18.” Oral abstract presented at The 20th Annual Meeting of the RNA Society, Madison, WI, USA. (2015 May).

“The Yeast Rpl22 Protein Inhibits the Splicing of RPL22B Pre-mRNA to Modulate the Levels of Mature Transcript.” Poster session presented at the Annual Retreat of the UCLA Molecular Biology Institute, Los Angeles, CA. (2015 April).

Chapter 1

An Overview of pre-mRNA Splicing, Splicing Autoregulation, and Alternative Splicing in *Saccharomyces cerevisiae*

Since the initial discovery of introns within the genome of adenovirus 2 in the late 1970s (Berget et al., 1977; Chow et al., 1977a, 1977b), tremendous progress has been made in understanding the roles of intron splicing within the overall process of gene expression. Spliceosomal introns have been found in all known eukaryotic genomes and display an astounding diversity in size and frequency that varies both within a given genome and when comparing genomes of different organisms (Irimia and Roy, 2014). Decades of study have uncovered numerous roles for introns, including their ability to modulate gene expression by impacting transcription rates, mRNA export, and translation efficiency (Shaul, 2017), as well as promoting genetic stability by limiting the formation of R-loops (Bonnet et al., 2017). Perhaps the most widely recognized function of introns is their contribution toward increasing protein diversity through the alternative splicing of intron-containing genes (Nilsen and Graveley, 2010). Illustrating this point, the presence of an average of 8 introns within each of the ~26,000 protein-coding genes in the human genome allows for the production of nearly four times that many unique proteins through regulated alternative splicing (Ast, 2004). The generation of multiple protein isoforms through pre-mRNA splicing necessitates a high degree of regulation and quality control, since the correct splicing of introns to form appropriately functioning proteins is essential to normal cellular operations. In humans, splicing errors contribute toward a number of systemic diseases and over 50% of human genetic disorders (Scotti and Swanson, 2015; Wang and Cooper, 2007). Moreover, RNA splicing is an essential function that serves as a major avenue for the regulation of gene expression in eukaryotes.

Overview of the Mechanism of Splicing by the Major Spliceosome

Splicing of the majority of introns in eukaryotes is accomplished though the actions of the major nuclear spliceosome, a large molecular complex consisting of five small nuclear ribonucleoproteins (snRNPs), U1, U2, U4, U5 and U6, as well as a large number of associated splicing factors that participate in specific stages of splicing (Bentley, 2014; Chen and Cheng,

2012; Matera and Wang, 2014; Nilsen, 1994). The spliceosomal snRNPs contain multiple constituent proteins and snRNAs that function together to facilitate splicing. RNA splicing depends on three major signals within an intron: the 5'-splice site (5'-SS), the branchpoint, and the 3'-splice site (3'-SS). The 3'-SS is usually preceded by a polypyrimidine tract that serves as a binding platform for specific proteins and promotes the assembly of the spliceosome (Roecigno et al., 1993). In general, the 5'-SS is characterized by a highly conserved GU motif at its 5' end, while the 3'-SS usually ends with an AG dinucleotide (Burset et al., 2000). The branchpoint, which is likewise part of a consensus sequence, is most often an adenosine that becomes bulged when the branchpoint sequence is hybridized with the U2 snRNA (Gao et al., 2008). In *S. cerevisiae*, the 5'- and 3'-SS have the consensus sequences of GUAUGU and UAG, respectively, while the consensus branch sequence is UACUAAC, in which the penultimate base is the branchpoint adenosine (Spingola et al., 1999). These three sites on the intron serve as the major handles of substrate recognition and catalysis for the spliceosome during RNA splicing (Fig. 1.1A).

Splicing occurs in a concerted manner concurrent with or immediately following transcription (Tilgner et al., 2012) beginning with the initial assembly of the spliceosome onto the pre-mRNA substrate. The U1 snRNP recognizes and binds to the 5'-SS while the branchpoint is bound by the branchpoint bridging/binding protein (BPP, SF1 in humans) (Berglund et al., 1997; Konarska, 1998). The polypyrimidine tract and 3'-SS are recognized and bound by Mud2 (U2AF in mammals), and the interaction of all of these bound components forms the commitment complex (Abovich and Rosbash, 1997; Rain and Legrain, 1997). This arrangement commits the pre-mRNA to the splicing reaction by bridging the splice sites and branchpoint together for subsequent chemical reactions (Fig. 1.1).

Assembly continues with the hydrolysis of ATP by the splicing factor Prp5, resulting in the exchange of BBP for the U2 snRNP in the binding of the branch sequence and a conformational change that generates the A complex (Shao et al., 2012). The imperfect complementarity between

the U2 snRNA and the branch sequence results in the branchpoint adenosine being “bulged” out of the base pairing interactions (Berglund et al., 2001). Next, the U4/U5/U6 tri-snRNP is recruited to the spliceosome in an ATP-dependent process catalyzed by Prp28, forming the B complex in which all five spliceosomal snRNAs are present in the spliceosome (Strauss and Guthrie, 1994). Final rearrangements are catalyzed by the RNA helicase Brr2 and the ATPase Prp2, which disrupt the interaction between the U4 and U6 snRNPs and release the U1 and U4 snRNPs from the spliceosome (Ohrt et al., 2012; Raghunathan and Guthrie, 1998). In the resulting conformational change, U6 replaces U1 in binding to the 5’-SS, while U2 binds with U6. The U5 snRNP tethers the two exons in order to position them for ligation in a later stage (Newman, 1997). During this state a large complex of proteins known as the nineteen complex (NTC, after the essential factor Prp19) also joins the spliceosome, where it stabilizes the binding of the U5 and U6 snRNPs within the spliceosome (Chan and Cheng, 2005; Chan et al., 2003). The spliceosome is now catalytically competent and is referred to as the B* complex (Fig. 1.1B).

The first catalytic step of splicing is an ATP-independent transesterification reaction in which the 2’-hydroxyl of the bulged adenosine of the intron branchpoint makes a nucleophilic attack on the 5’-SS, cleaving the first exon from the intron and forming a 2’-5’ phosphodiester linkage between the 5’-SS G_1 guanosine and the branchpoint adenosine, thereby generating an intron lariat-exon 2 RNA intermediate and rearranging the spliceosome into the C complex (Fig. 1.1). This reaction requires the activities of the splicing factors Yju2 and Cwc25 (Chiu et al., 2009; Liu et al., 2007). Although the first catalytic step does not directly require ATP hydrolysis, Yju2 and Cwc25 are only able to promote the reaction after the ATP-dependent spliceosomal rearrangements catalyzed by Prp2 are completed.

The U5 snRNP is central to the second catalytic step of splicing. The U5 snRNA contains an invariant loop of nine nucleotides that interact with the exonic nucleotides of the RNA substrate that flank the intron, positioning them for ligation within the catalytic core of the spliceosome

(Frank et al., 1994; Newman and Norman, 1992). Remarkably, while these exonic sequences can determine the stability of the interaction with loop 1 nucleotides of the U5 snRNA and thus influence the fidelity 3'-SS selection (Crotti and Horowitz, 2009), they are poorly conserved in yeast (Horowitz, 2012), suggesting that 3'-SS choice may be more dependent upon other determinants.

Following the first catalytic reaction, hydrolysis of ATP by Prp16 remodels the spliceosome and releases Yju2 and Cwc25 to form the C* complex (Fig. 1.1) (Tseng et al., 2011). The timing of Prp16 function is highly significant – if Prp16 binds the spliceosome before the first transesterification reaction has occurred, it may prematurely hydrolyze ATP and eject Cwc25 and Yju2, effectively rejecting the pre-mRNA substrate (Koodathingal et al., 2010; Tseng et al., 2011). Because slow splicing kinetics are often dictated by poor splice site choice, this Prp16-mediated rejection of slow substrates is believed to contribute toward the overall fidelity of splicing (Horowitz, 2011).

Following the release of Prp16, Yju2, and Cwc25, the second-step splicing factors Slu7 and Prp18 bind the spliceosome and associate with the U5 snRNP (Fig. 1.1) (James et al., 2002). Prp18 and Slu7 interact both functionally and physically to position the 3'-SS within the active site of the spliceosome (Ohrt et al., 2013). Furthermore, both proteins have been shown to impact 3'-SS choice (Frank and Guthrie, 1992; Kawashima et al., 2014), and both are dispensable for splicing when the distance between the 3'-SS and the branchpoint is sufficiently short (Brys and Schwer, 1996; Zhang and Schwer, 1997).

The second transesterification reaction involves a nucleophilic attack by the 3'-hydroxyl at the end of exon 1 on the 3'-SS, forming a 5'-3' phosphodiester link and resulting in the cleavage of the 3'-SS from exon 2, the separation of the intron lariat from the exons, and the joining of the two exons (Fig. 1.1). Similar to the first catalytic step, the second catalytic reaction is ATP-independent (Horowitz and Abelson, 1993). Release of the mRNA from the post-spliceosomal

complex involves the translocation and helicase activities of Prp22, which joins the spliceosome prior to the second transesterification but after the binding of Slu7 and Prp18. The translocation of Prp22 along the RNA requires ATP hydrolysis and results in the ejection of Slu7 and Prp18 from the spliceosome, while the intrinsic helicase activity severs the interactions between the U5 snRNA loop 1 nucleotides and the exonic nucleotides of the mRNA, thus releasing the spliced mRNA and forming the post-spliceosomal complex (Schwer, 2008; Schwer and Meszaros, 2000). Mutants of Prp22 have also been shown to allow exon ligation for pre-mRNAs harboring 3'-SS sequence mutations that would otherwise inhibit that step, indicating that Prp22 normally proofreads exon ligation (Mayas et al., 2006).

The final step of splicing involves the removal of the remaining U2/U5/U6 snRNPs and associated factors from the intron lariat to which they remain bound following the second transesterification reaction. The removal of these snRNPs is accomplished through the binding of Ntr1, Ntr2, and Prp43, after which the helicase activity of Prp43 removes the remaining snRNPs to be recycled for another splicing cycle (Fig. 1.1) (Arenas and Abelson, 1997; Fourmann et al., 2017; Martin et al., 2002; Tsai et al., 2005). Interestingly, Prp43 has also been shown to cooperatively proofread 5'-splice site cleavage by disposing of suboptimal pre-mRNA substrates that are rejected by Prp16 (Koodathingal et al., 2010). The newly spliced messenger ribonucleoprotein (mRNP) is prepared for export to the cytosol through to the combined efforts of multiple mRNA export factors, including the essential proteins Mex67 and Mtr2 (Stewart, 2010). Meanwhile, the phosphodiester bond within the intron lariat released from the spliceosome is cleaved by Dbr1 (Chapman and Boeke, 1991) and the linearized intron is degraded by nuclear exonucleases (Hesselberth, 2013; Moore, 2002).

Regulation of Splicing in *Saccharomyces cerevisiae*

The budding yeast *Saccharomyces cerevisiae* has proven an invaluable organism for the study of RNA splicing, owing partly to the simplicity of its intronome. Of the ~6000 genes within

the *S. cerevisiae* genome, only about 3-4% contain introns and almost all of these have only a single intron (Ast, 2004; Lin et al., 2013). Furthermore, the majority of yeast introns can be removed from the genome with little impact on cell fitness under normal growth conditions (Parenteau et al., 2008), prompting speculation as to the necessity of introns in this organism. However, in spite of their relative scarcity, intron-containing genes in *S. cerevisiae* supply roughly 27% of all mRNAs synthesized in the cell, suggesting that introns nevertheless play a critical role in the expression of yeast genes and consequently in overall cellular functionality (Ares et al., 1999). The relative dearth of introns in the *S. cerevisiae* genome has traditionally led to the generalization that the regulation of splicing is primarily achieved through the regulation of intron retention. However, the simplicity of the yeast intronome and the highly conserved nature of the core spliceosome components throughout Eukaryota (Jurica and Moore, 2003) provides a useful backdrop to study the intricacies of splicing and the surprising multitude of players that determine its efficiency and progression, including variability in substrate recognition, the propensity to form pre-mRNA secondary structures that can sequester or expose splicing signals, and the involvement of chromatin structure in the regulation of splicing, to name a few (reviewed in Johnson and Vilardell, 2012). As will be explored in detail in subsequent chapters of this dissertation, the regulation of both canonical and alternative splicing constitute major avenues for gene expression control in yeast. The remainder of this introduction will elaborate upon two of these regulatory schemes that have been explored in *S. cerevisiae*: autoregulation of the splicing of certain pre-mRNAs and functional and non-productive alternative splicing. Subsequent chapters will highlight specific studies addressing key questions involving these two classes of regulation.

Intron-mediated Autoregulation of Gene Expression in *S. cerevisiae*

Studies in pre-mRNA splicing across multiple eukaryotes have revealed numerous examples of splicing-mediated autoregulation, in which the protein product of a gene influences

the splicing of its own pre-mRNA transcript *in trans*. In *S. cerevisiae*, nearly 50% of intron-containing genes encode ribosomal proteins (RPs) and ribosomal protein gene (RPG) transcripts constitute the largest proportion of cellular mRNAs (Ares et al., 1999). It is therefore not surprising that several cases of autoregulation involve genes coding for ribosomal proteins, exemplifying extra-ribosomal functionality beyond the translational roles of the ribosome (Lu et al., 2015; Warner and McIntosh, 2009). Arguably the earliest example of splicing-mediated autoregulation in yeast was identified for the protein Rpl30, whose pre-mRNA transcript folds into a secondary structure that base pairs nucleotides from exon 1 with those of the 5'-splice site (Eng and Warner, 1991). This arrangement is further stabilized by the binding of the Rpl30 protein to the secondary structure (Fig. 1.2A). Although the Rpl30-mediated sequestration of the 5'-splice site was originally hypothesized to prevent the U1 snRNP from accessing this critical site of splicing chemistry, further analysis of the Rpl30-mRNA complex revealed that the U1 snRNP is able to assemble normally and that the Rpl30 protein instead prevents the U2 snRNP from assembling onto the nascent spliceosome (Fig. 1.2A) (Chao and Williamson, 2004; Vilardell and Warner, 1994).

The aforementioned case of Rpl30 involves the regulation of a single-locus RPG by its protein product. However, in *S. cerevisiae*, 59 ribosomal protein genes have paralogs that arose from a whole genome duplication event (Steffen et al., 2012). Although several of these paralogs encode identical or very similar proteins, the expression and regulation of paralogous RPGs have nevertheless been shown to be asymmetrical (Parenteau et al., 2015), giving rise to the designations of major and minor paralogs. Indeed, autoregulatory mechanisms for duplicated intron-containing RPGs have been found to differentially affect paralogous transcripts. In a recent study, the pre-mRNAs for *RPS9A* and *RPS9B*, both of which code for the ribosomal protein Rps9, were shown to possess dissimilar intron structures and 3'UTR lengths that impact their splicing efficiencies (Petibon et al., 2016). Specifically, the minor paralog *RPS9A* contains an intronic

helical structure that suppresses its splicing via recognition by the Rps9 protein, where the degree of splicing inhibition is exacerbated with increasing concentrations of the Rps9 protein. Although the pre-mRNA of the major paralog *RPS9B* also contains an intron, its affinity for Rps9 is substantially lower, resulting in a lack of splicing inhibition. Thus, the Rpl9 protein preferentially affects the processing of one of its paralogous transcripts over the other. Similarly, a study of the yeast RPG *RPL22* revealed that the pre-mRNA of the minor paralog *RPL22B* contains an intronic stem loop structure that is recognized and bound by the Rpl22 protein (Gabunilas and Chanfreau, 2016), see Chapter 3). The binding of Rpl22 to the regulatory element significantly inhibits the splicing of the *RPL22B* pre-mRNA (Fig. 1.2B). It was further shown that this autoregulatory mechanism is important for downregulating the levels of mature *RPL22B* transcripts in DNA damage stress (Gabunilas and Chanfreau, 2016) perhaps to allow the limited supply of spliceosomal components to access other transcripts under these conditions (Munding et al., 2013).

Interestingly, autoregulatory circuits for Rps9 and Rpl22 were found to be conserved in higher eukaryotes. Overexpression of the RpS9 protein in *Drosophila melanogaster* was shown to drive the production of alternatively spliced mRNA isoforms of chromosomally-encoded *RPS9* that are subsequently degraded by the nonsense-mediated mRNA decay (NMD) pathway, a conserved cytoplasmic translation-dependent RNA degradation pathway that identifies transcripts containing premature termination codons (Kervestin and Jacobson, 2012; Lykke-Andersen and Jensen, 2015). Thus, RpS9 autoregulation serves as a means to fine tune overall *RPS9* transcript levels (Plocik and Guthrie, 2012). Likewise, autoregulatory pathways for the expression of orthologous Rpl22 proteins in zebrafish and mice similarly depend on the recognition of RNA transcript secondary structures in the paralogous transcript that are specifically recognized by Rpl22 (O'Leary et al., 2013). While it is unclear whether the metazoan Rpl22 regulatory mechanisms are dependent upon the suppression of pre-mRNA splicing as in yeast, the

preservation of autoregulatory pathways for this protein in multiple eukaryotes may nevertheless indicate either a preserved functional relevance that has persisted throughout evolution or, alternatively, an emergent function resulting from convergent evolution.

Finally, large-scale deletion analyses of RPG introns in *S. cerevisiae* have revealed their central role in regulating the expression of their respective genes and their paralogs (Parenteau et al., 2011). As discussed above, this intron-dependent regulation oftentimes differentially impacts paralogous RPGs. Further studies will be needed to uncover the molecular mechanisms underlying these intra- and inter-paralog regulatory effects for each unique set of genes, to understand why paralogous RPGs have been maintained throughout evolution, and more generally to elucidate the need for enriching one of the two duplicated RPG paralogs. To the latter point, one potential reason for maintaining slightly different copies of the same gene might be to modulate ribosome function according to cellular stresses, for instance to selectively translate certain transcripts. A unique case of this type of ribosome reprogramming in yeast is exemplified by the enrichment of Rpl22A over the paralog Rpl22B in functional ribosomes in response to oxidative stress (Chan et al., 2012). This reprogramming was predicted to promote the selective translation of certain mRNAs by the ribosome, although the full impact of the changes in ribosome composition on the cellular protein repertoire remains to be determined.

Intron-mediated autoregulatory feedback mechanisms for intron-containing genes have also been elucidated for non-RPG transcripts. A remarkable example in *S. cerevisiae* was demonstrated for the DEAD-box RNA helicase Dbp2, which has an usually long 3'-proximal intron (Barta and Iggo, 1995). In their study, Barta and Iggo found that the deletion of the *DBP2* intron dramatically increased the levels of the Dbp2 protein and that the overexpression of exogenous Dbp2 protein suppresses the levels of Dbp2 protein expressed from the endogenous chromosomal allele. They further showed that large fragments of the *DBP2* intron that were transposed into foreign genes were able to confer the suppressive effect by the Dbp2 protein onto

the chimeric transcripts, indicating a negative feedback loop for Dbp2 that relies upon intron recognition (Barta and Iggo, 1995). Subsequent studies revealed that Dbp2 interacts with the RNA binding protein Yra1 during nuclear messenger ribonucleoprotein (mRNP) assembly and that this interaction inhibits the helicase activity of Dbp2 (Ma et al., 2013). Strikingly, a splicing-mediated autoregulatory mechanism has also been identified for *YRA1* that relies on two distinct regulatory elements located within exon 1 and the intron of the *YRA1* pre-mRNA (Dong et al., 2007). Specifically, the exon 1 element is recognized by the Yra1 protein which inhibits splicing of the pre-mRNA while the intronic regulatory element is necessary for the cytoplasmic degradation of the pre-mRNA, the latter of which requires the mRNA export factor Mex67 and the decapping protein Edc3 (Dong et al., 2007). While it is possible that the RNA-based autoregulatory mechanisms for the interacting factors *DBP2* and *YRA1* evolved independently of one-another, it is tempting to speculate that they may serve a collaborative functional purpose in regulating the assembly of mRNPs or modulating the nuclear export of pre-mRNAs, perhaps under conditions in which rebalancing between pre-mRNA splicing and mRNA export must be achieved (Johnson and Vilardell, 2012).

Recent studies utilizing RNA structure prediction tools have also uncovered novel non-coding RNA (ncRNA) structures contained within the introns of several transcripts in *S. cerevisiae* that were shown to modulate the expression of their host genes, in some cases in a stress-dependent manner (Hooks et al., 2016). Notably, ncRNA structures were shown to promote intron retention in the RPG-coding pre-mRNA transcripts that contain them, relative to intron-containing RPGs with no predicted ncRNAs, reminiscent of the intron structure-dependent regulation of splicing for *RPL30* and *RPL22B* (Fig. 1.2) (Eng and Warner, 1991; Gabunilas and Chanfreau, 2016; Hooks et al., 2016). Future investigations are likely to elucidate the molecular mechanisms governing these and other regulations for intron-containing genes in yeast.

Alternative Splicing in *Saccharomyces cerevisiae*

Owing to the general scarcity of introns within the *S. cerevisiae* genome, the traditional viewpoint has held that alternative splicing does not occur in budding yeast for the classical purpose of increasing the diversity of the proteome. While it has been known for some time that changes in cellular environment or the introduction of cellular stresses can trigger substantial changes in intron retention (Bergkessel et al., 2011; Pleiss et al., 2007), the scope of alternative splicing with regards to the ability to generate multiple spliced forms of a transcript have until recently gone underappreciated. Indeed, studies have shown that RNA splicing in yeast encompasses numerous cases of functional alternative splicing, exon skipping, and non-productive alternative splicing for the regulation of gene expression, suggesting that the impact of splicing extends beyond the creation of translation-competent transcripts. The following sections will describe some of these cases in greater detail.

Functional Alternative Splicing

In higher eukaryotes, protein diversity can be achieved through the selective inclusion or exclusion of certain exons to generate multiple transcript and protein isoforms from a single genetic locus (Black, 2003; Kornblihtt et al., 2013). By contrast, intron-containing genes in *S. cerevisiae* have only one or (rarely) two annotated introns, leading to the general conclusion that the yeast proteome does not experience diversification in this manner. Nevertheless, there exist documented examples of alternative splicing in budding yeast that give rise to functional proteins.

One such example involves the Hsp70 nucleotide exchange factor (NEF) Fes1, a key component in protein quality control that promotes the degradation of misfolded proteins by the proteasome (Gowda et al., 2013). The *FES1* gene contains a cryptic intron beginning at the 3' end of the coding sequence that bridges exon 1 with downstream nucleotides that constitute a second exon (Gowda et al., 2016). When left unspliced, the *FES1* transcript is polyadenylated just downstream of exon 1, and the resulting mRNA is translated as a short protein isoform Fes1S (Fig 1.3). However, the removal of the intron excises the regular polyadenylation site, extends the

coding sequence by several nucleotides and shifts polyadenylation to a distal site downstream of exon 2. The translation of this alternatively spliced *FES1* transcript results in a long protein isoform Fes1L (Fig. 1.3) (Gowda et al., 2016). Thus, the alternative splice isoforms of *FES1* are the result of competition between the proximal polyadenylation and splicing of the intron. While both isoforms possess NEF activity, Fes1S production is enhanced in heat shock conditions as a result of heat shock-regulated alternative splicing. Fes1S is cytosolically localized, is necessary for the degradation of misfolded proteins via downstream ubiquitin ligases, and tempers the heat shock response in unstressed conditions. Conversely, Fes1L is expressed at comparatively lower levels than Fes1S both in stressed and unstressed conditions and is targeted to the nucleus via a nuclear localization signal (NLS) (Fig. 1.3). Surprisingly, Fes1L is not necessary for the degradation of misfolded proteins in the nucleus and can compensate for the absence of Fes1S when overexpressed (Gowda et al., 2016). The persistence of the Fes1L isoform suggests a physiological function, but the biochemical role of this protein has yet to be conclusively identified.

A well-known example of functional alternative splicing in *S. cerevisiae* has been described for the transcript encoding the protein phosphatase Ptc7, a mitochondrial protein responsible for activating the biosynthesis of coenzyme Q6 via the SWI/SNF complex in response to metabolic shifts (Awad et al., 2017; Jiang et al., 2002; Martín-Montalvo et al., 2013). The *PTC7* transcript contains a 5'-proximal intron that, when left unspliced, does not disrupt the reading frame and contains no in-frame stop codons (Nash et al., 2007). Consequently, both the spliced and unspliced forms of the transcript encode fully functional proteins with apparently distinct physiological roles (Juneau et al., 2009). The spliced form of the transcript gives rise to the $Ptc7_s$ protein, which localizes to the mitochondria where it fulfills the documented function in coenzyme 6 biosynthesis. By contrast, the unspliced *PTC7* transcript includes coding sequences for a putative transmembrane domain, and the resulting protein $Ptc7_u$ localizes to the nuclear envelope. In accordance with its role in cellular respiration, the levels of $Ptc7_s$ protein were notably enhanced

in cells that were grown on non-fermentable carbon sources, although this increase was not due to changes in *PTC7* pre-mRNA splicing (Juneau et al., 2009). $Ptc7_u$ appears to fulfill an entirely different function – its expression was shown to mitigate the deleterious effects of treating cells with Latrunculin A, a chemical disruptor of actin filament formation. Therefore, it may potentially play a role in the formation or stability of the cytoskeleton. Remarkably, comparative sequence analysis between the orthologous *PTC7* genes of over a dozen yeast species not only revealed a high level of conservation of intronic splicing signals, but also the presence of putative transmembrane domains and mitochondrial targeting signals for long and short orthologous Ptc7 protein isoforms (Juneau et al., 2009). This finding indicates that, at least among yeast, the ability to generate distinct protein isoforms from the *PTC7* locus serves an important functional purpose.

The aforementioned examples of *FES1* and *PTC7* represent two of the very few known examples of functional alternative splicing in budding yeast. However, it is possible that at one point over the course of evolution the production of multiple distinct proteins from a single genetic locus through alternative splicing was a more common phenomenon in yeast. For instance, in many fungal species the distinct proteins Ski7 and Hbs1 are encoded by a single alternatively spliced gene (Marshall et al., 2013). In *S. cerevisiae*, however, the *SKI7* and *HBS1* genes that resulted from whole genome duplication have since become subfunctionalized – their transcripts are no longer spliced and their protein products have distinct roles in RNA metabolism. Specifically, Ski7 links the Ski complex with the cytoplasmic exosome to mediate the 3'-to-5' exonucleolytic degradation of RNA (Houseley and Tollervey, 2009), while Hbs1 works with its binding partner Dom34 to dissociate stalled ribosomes from mRNAs as part of the no-go decay mRNA degradation pathway (Chen et al., 2010; Tsuboi et al., 2012). The evolutionary rationale for the subfunctionalization of distinct genes versus protein diversification by alternative splicing remains a topic of great interest.

Regulation of Two-Intron Genes

A very small minority of yeast genes contain two annotated introns whose presence expands the avenues for gene expression control. For two-intron genes, much of this control is exercised by way of RNA turnover. An analysis of the splicing and degradation patterns for several two-intron transcripts revealed an important role in the regulation of these transcripts by nuclear RNA surveillance and turnover pathways (Egecioglu et al., 2012). The transcriptional regulator *MAT*a1, a homeobox-domain protein that regulates the expression of haploid-specific genes (Mathias et al., 2004), showed a substantial accumulation of splicing intermediates in yeast strains mutated for the nuclear endonuclease Rnt1, the 5'-to-3' exonuclease Rat1, or the 3'-to-5' nuclear exosome component Rrp6. Specifically, the unspliced and partially spliced (retaining intron 2) *MAT*a1 transcripts are able to form a stem-loop secondary structure that is recognized and cleaved by Rnt1, with the cleavage products being subsequently degraded by nuclear exonucleases (Fig. 1.4). In the absence of Rnt1 cleavage, the unspliced and partially spliced transcripts can either be correctly spliced, exported and expressed as the 'a1' protein, or alternatively spliced to form an exon2-skipped isoform in which exon 1 is directly ligated to exon 3 (Egecioglu et al., 2012). The exon2-skipped transcript is then degraded by Rat1. The interplay between productive and non-productive splicing of *MAT*a1 by the spliceosome, cleavage by Rnt1, and degradation by the nuclear RNA turnover machinery allows for intricate control of *MAT*a1 expression and prevents the translation of aberrant transcripts into truncated proteins (Fig. 1.4). Quality control of exon skipping events was also demonstrated for the two-intron transcripts *SUS1*, *DYN2*, and *YOS1*, in which exon2-skipped isoforms for all three transcripts accumulated in mutants for Rat1 and Rrp6 (Egecioglu et al., 2012). These results highlight the role of nuclear RNA surveillance pathways in suppressing expression of incompletely spliced and exon-skipped species.

Independent investigations of *DYN2* and *SUS1* have revealed additional insights into the regulation of multi-intron genes. *DYN2* was utilized as a case study to evaluate the impact of *cis*

and *trans*-acting determinants on the inclusion of the internal exon (Howe et al., 2003). *Cis* mutations of the branchpoint sequence of intron 1 and a mutation in the U2 snRNA *in trans* were shown to promote skipping of the internal exon, demonstrating that the efficient recognition of the branchpoint sequence by the spliceosome is critical for the inclusion of that exon. Exon inclusion was also shown to be heavily dependent upon the rate of transcription elongation – mutants of the RNA polymerase II (Pol II) subunit *RBP2* that exhibit slowed elongation rates shifted *DYN2* splicing toward greater exon 2 inclusion (Howe et al., 2003). Similarly, the treatment of yeast cells with drugs that inhibit nucleotide biosynthesis and therefore enhance the frequency of Pol II pausing in conditions of reduced nucleotide availability was also shown to suppress exon skipping (Howe et al., 2003). These observations provided additional support for the first come, first served model, in which the splicing signals that are recognized earlier or more efficiently are the ones that are most readily utilized by the spliceosome. In the case of *DYN2*, suboptimal recognition of the intron 1 branchpoint sequence allows Pol II to transcribe the downstream intron 2 before the splicing of intron 1 is initiated. The branchpoint sequence of intron 2 may then be paired with the 5'-SS of intron 1 for the first transesterification reaction, resulting in exon 2 being included in the intron lariat. Slow Pol II mutants allow more time for intron 1 to be spliced correctly without competition from splicing signals in the second intron, thereby suppressing exon skipping (Howe et al., 2003). This study showcases the impact of the co-transcriptional nature of splicing on determining how the nascent transcript is ultimately processed.

The mRNA export factor Sus1 plays an important role in histone H2B deubiquitination and is encoded by a two-intron gene whose splicing was shown to be influenced by the mRNA cap binding complex (Hossain et al., 2009; Köhler et al., 2006). The first intron of the *SUS1* pre-mRNA contains non-consensus, sub-optimal 5'-SS and branchpoint sequences, whereas the branchpoint and 5'SS sequences of the second intron are consensus. The non-consensus nature of the intron 1 splicing signals contributes towards its partial retention during splicing that is

exacerbated in conditions of environmental stress (Hossain et al., 2011). The inclusion of intron 1 introduces a premature termination codon (PTC) into the reading frame of *SUS1*, and consequently the intron 1-retained transcript is a substrate for NMD, a property shared by many intron-retained transcripts (Sayani et al., 2008). Interestingly, deletion of the first intron from the *SUS1* gene led to a slight increase in the retention of the other intron, suggesting that the assembly of the splicing machinery on intron 1 influences the processing of the second intron (Hossain et al., 2011). Surprisingly, genetic complementation of a temperature-sensitive, H2B deubiquitination-deficient *sus1*Δ null mutant strain with the intron-containing wild-type *SUS1* gene, but not with intronless *SUS1* cDNA, was able to suppress the physiological defects of that strain. Thus, the ability to generate partially- or alternatively-spliced *SUS1* mRNA isoforms that presumably encode distinct protein products is important for cellular functionality (Hossain et al., 2011). Alternatively, the incomplete splicing of *SUS1* could serve as a means to regulate the levels of the *SUS1* mRNA and consequently of the Sus1 protein depending on the environmental conditions. Altogether, the regulation of *SUS1* splicing not only illustrates the intragenic effects of introns in the processing of the pre-mRNA, but also demonstrates the importance of introns in maximizing cellular fitness.

Non-productive Alternative Splicing

Splicing has a direct impact on the translatability of intron-containing protein-coding genes. With few exceptions (e.g. *PTC7*, see above), retained introns have a high probability of introducing PTCs into the reading frame, resulting in nonsense transcripts that could potentially encode nonfunctional or malfunctioning proteins with deleterious biological effects. Although intron retention is widespread throughout the yeast transcriptome, these events frequently go unobserved due to the degradation of these transcripts by NMD (Sayani et al., 2008). Additionally, several pre-mRNA transcripts harboring retained introns are subject to degradation by both the nuclear exosome and NMD and are only stabilized when both of these pathways are disabled,

indicating that the nuclear exosome and by NMD are sequential pathways that serve to limit the accumulation of unspliced RNA (Sayani and Chanfreau, 2012). Therefore, intron retention directly affects the pool of translation-competent transcripts.

In addition to retained introns, PTCs may also be introduced into protein-coding transcripts via splicing events at cryptic 5'- and 3'-SS that shift the translational reading frame (Fig. 1.5). In contrast to the cases of functional alternative splicing described above, non-productive alternative splicing yields products that are targeted by cellular RNA degradation mechanisms, most notably NMD (Fig. 1.5). As such, many cryptic splicing events are undetectable in the presence of the functional NMD pathway. Indeed, a study of a wide panel of yeast splicing factor mutants showed that several of these mutations promote cryptic splicing or the accumulation of unspliced transcripts, but that these events are masked by NMD (Kawashima et al., 2009). This finding suggested that the inactivation of NMD is necessary in order to fully understand the roles of individual splicing factors in maintaining splicing efficiency and fidelity (see Chapter 4). A follow-up investigation analyzing the transcriptome of NMD-compromised yeast by next-generation RNA-seq found that the use of cryptic, non-productive alternative splice sites is widespread in *S. cerevisiae* and that PTC-containing transcripts resulting from alternative splicing are efficiently degraded by NMD ((Kawashima et al., 2014), see Chapter 2). This implies that cryptic alternative splicing coupled to NMD (AS-NMD) serves to regulate transcript levels in yeast, as opposed to alternative splicing that promotes the diversification of the proteome. In one specific case, alternative splicing of the *RPL22B* transcript was also shown to be promoter-dependent and responsive to particular stress conditions, demonstrating the cell's utilization of alternative splicing to regulate transcript levels in response to environmental changes (Kawashima et al., 2014).

Similar RNA-seq analyses have identified additional novel cryptic splicing events and instances of intron retention in yeast transcripts (Schreiber et al., 2015), highlighting the utility of RNA-seq in the detection of low-abundance splicing events and transcript variants. Interestingly,

AS-NMD has also been observed in plants and across metazoans as a means to regulate the expression of specific protein isoforms (de Lima Morais and Harrison, 2010; Smith and Baker, 2015). In fact, several SR and hnRNP genes that encode splicing factors in humans and mice are targets of AS-NMD, suggesting that the splicing mechanism itself may be subject to negative feedback suppression (McGlincy and Smith, 2008). Finally, multiple human cancers and neurological disorders have been associated with the loss of regulation of human oncogenes, tumor suppressor genes, and neuronal transcripts through AS-NMD (da Costa et al., 2017). Thus, it is possible that AS-NMD originally evolved as a simple mechanism for the disposal of non-productive RNAs and was subsequently retained in higher eukaryotes where it may directly affect protein diversity, impact entire biochemical pathways, and influence systemic physiology.

Spliceosome-mediated Decay

Although non-productive with respect to protein biogenesis, some alternative splicing events are nevertheless functionally important from a physiological standpoint. Certain "intronless" yeast genes have been shown to harbor unannotated introns that, when retained, contribute toward the normal coding sequence. However, the removal of these unannotated introns by the spliceosome renders the transcript non-productive and subject to degradation. This process, termed spliceosome-mediated decay (SMD) (Fig. 1.6), was initially demonstrated in *S. cerevisiae* for the transcription initiation and chromatin remodeling factor *BDF2*, whose RNA transcript contains an unannotated intron with consensus splicing signals (Volanakis et al., 2013). Progression of splicing through the first catalytic step utilizing a 5'-SS proximal branchpoint sequence results in a cleavage of the *BDF2* mRNA at the 5'-SS, where the subsequent cleavage products are subject to degradation by the nuclear exosome. Alternatively, a two-step splicing pathway utilizing a downstream branchpoint sequence produces a fully spliced, truncated *BDF2* transcript that is likewise degraded by the nuclear exosome (Volanakis et al., 2013). Mutations of the *BDF2* intronic splicing signals or of particular splicing factors resulted in an increase in cellular

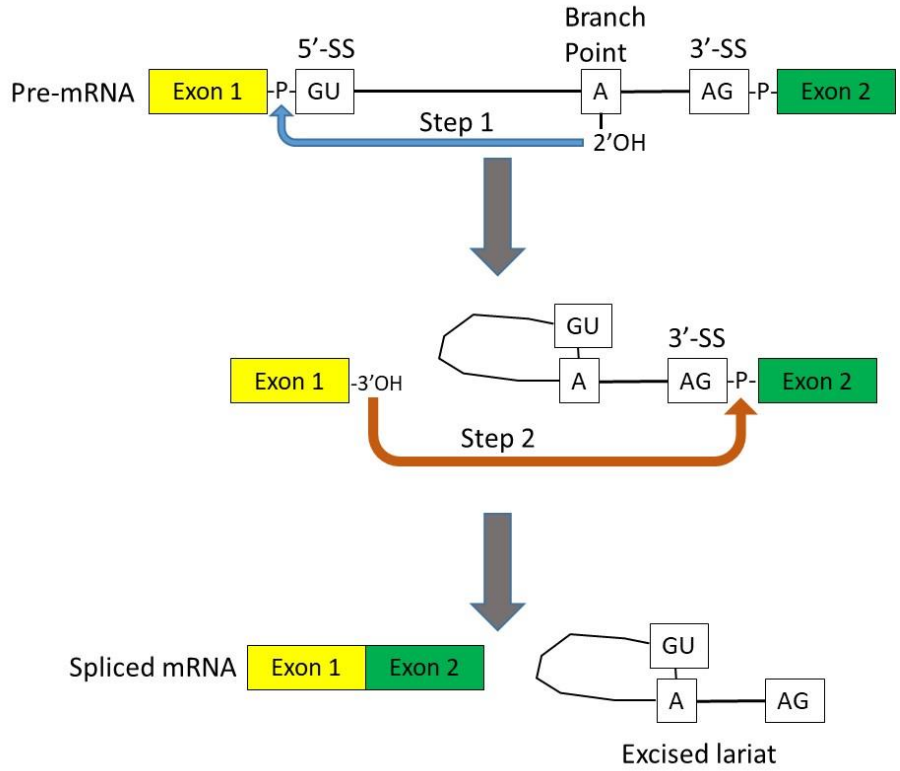
BDF2 mRNA levels, indicating that splicing directly impacts *BDF2* transcript abundance. Strikingly, the elevated levels of *BDF2* resulting from mutation of the 5'SS partially rescued the temperature sensitivity of mutant strains lacking the paralogous gene *BDF1*, hinting at the physiological importance of regulating Bdf2 protein levels (Volanakis et al., 2013). This work was expanded upon in a later study that revealed a stress-dependent dichotomy for the regulation of *BDF2* transcripts in the nucleus, wherein the degradation of *BDF2* by SMD is specifically inactivated in salt stress but is hyperactive during DNA damage stress (Roy and Chanfreau, 2014). Conversely, *BDF2* is predominantly cleaved by the Rnt1 endonuclease in salt stress, triggering the degradation of *BDF2* transcripts in a pathway termed Rnt1-mediated decay (RMD). In agreement with the earlier findings, disabling either SMD or RMD in *bdf1*Δ yeast cells subjected to stress conditions that hyperactivate those modes of *BDF2* transcript degradation rendered the *bdf1*Δ cells less sensitive to those respective stresses (Roy and Chanfreau, 2014). Taken together, these studies not only highlight the role of the spliceosome in the degradation of an RNA transcript through splicing of an unannotated intron (Fig 1.6), but also demonstrate that this degradation mechanism has a physiological impact on cellular fitness in specific stress conditions. SMD has been implicated in the degradation of other intronless transcripts as well (Volanakis et al., 2013), and it is possible that the regulation of these genes by SMD may also prove to be biologically important in a certain conditions of cellular stress.

Further Investigations into Splicing-associated Gene Regulation in Yeast

As illustrated by the multiple cases outlined the previous sections, the regulation of splicing in *S. cerevisiae* and its implications in gene expression extend far beyond the simple retention of introns. First, multiple yeast proteins are intricately involved in the processing of their own transcripts, including during pre-mRNA splicing, allowing for feedback-mediated regulation of gene expression. Furthermore, spliceosomes in budding yeast can utilize non-consensus alternative splicing signals, even in transcripts lacking annotated introns, to yield non-productive

transcripts that are targeted for degradation by RNA quality control pathways. Alternative splicing coupled to RNA decay can thereby limit the accumulation and ultimately the expression of the impacted transcripts. Additionally, because many alternative splicing and intron retention events are masked by the NMD pathway, analysis of RNA splicing patterns in NMD-compromised cells may unveil the roles of splicing factors in preserving the fidelity or promoting the efficiency of splicing. The following chapters will explore each of these points in greater depth by highlighting specific cases that have been investigated by our research group.

A.



B.

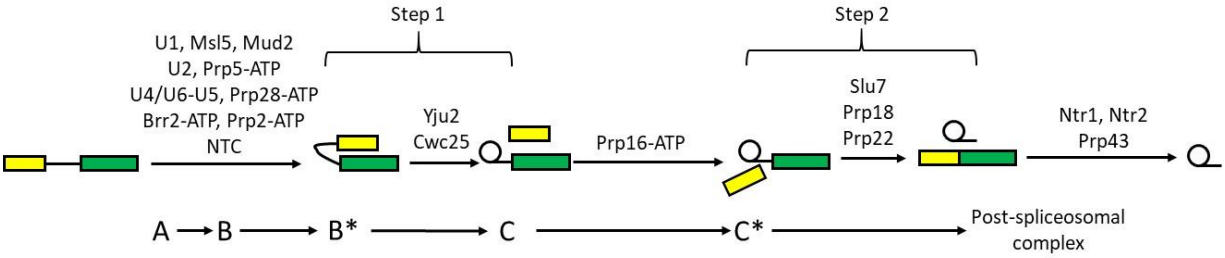


Figure 1.1. Simplified RNA-centric overview of the catalytic steps of splicing. A. The catalytic steps of pre-mRNA splicing involve two transesterification reactions. In the first reaction consists of a nucleophilic attack by the 2'-OH of the branch point adenosine on the 5'-phosphate of the 5'-SS guanosine. The second step involves another nucleophilic attack by the 3'-OH of Exon 1 on the phosphate group linking the 3'-SS to Exon 2. **B.** *Upper series:* Simplified overview of the incorporation of splicing factors into the spliceosome before, during, and after the catalytic stages. *Lower series*: Approximate stages of formation for the spliceosomal complexes (see text)

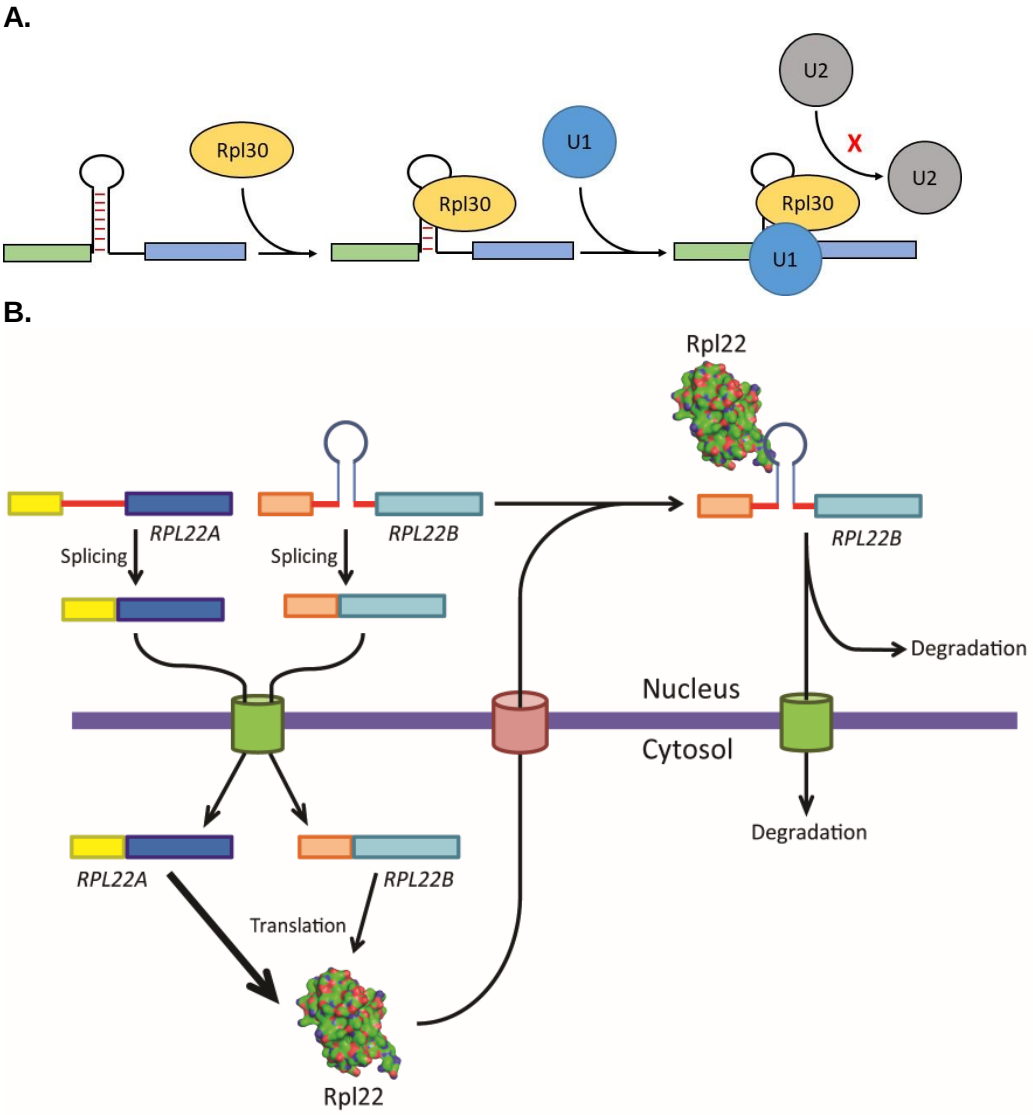


Figure 1.2. Intron-mediated autoregulation of Rpl30 and Rpl22. A. The *RPL30* pre-mRNA forms a secondary structure that sequesters the 5'-SS. This structure is stabilized by the binding of Rpl30 to the mRNP, which in turn prevents the association of the U2 snRNP, thereby inhibiting splicing. **B.** The pre-mRNA for *RPL22B* harbors an intronic stem loop regulatory element that is recognized and bound by the Rpl22 protein. The binding of the protein to the *RPL22B* pre-mRNA inhibits splicing of the transcript, which is then degraded by nuclear or cytoplasmic RNA degradation pathways.

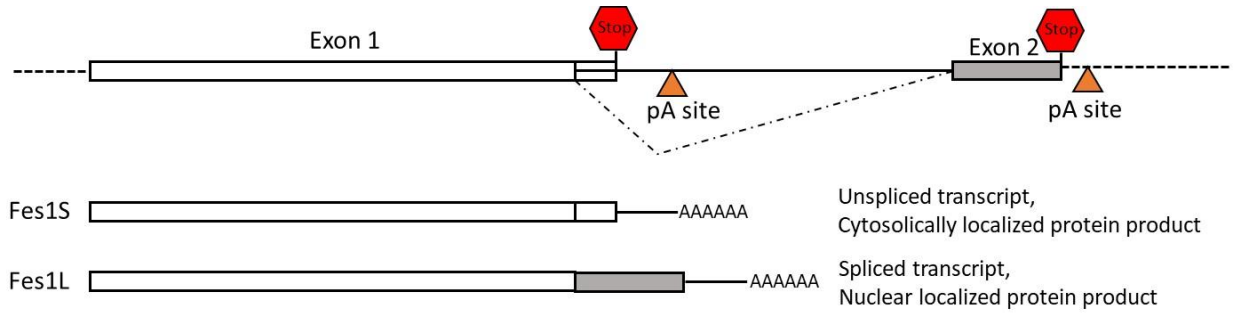


Figure 1.3. Splicing-dependent differential maturation of the *FES1* transcript. The *FES1* RNA transcript contains a cryptic intron that begins within the reading from of Exon 1. When left unspliced, *FES1* is polyadenylated at the upstream poly(A) site to give rise to the Fes1S isoform. If the intron is spliced out, the upstream poly(A) site is removed, the coding region is extended to include Exon 2, and the transcript is poly-adenylated at a downstream poly(A) site, producing the Fes1L isoform. Adapted from (Gowda et al., 2016).

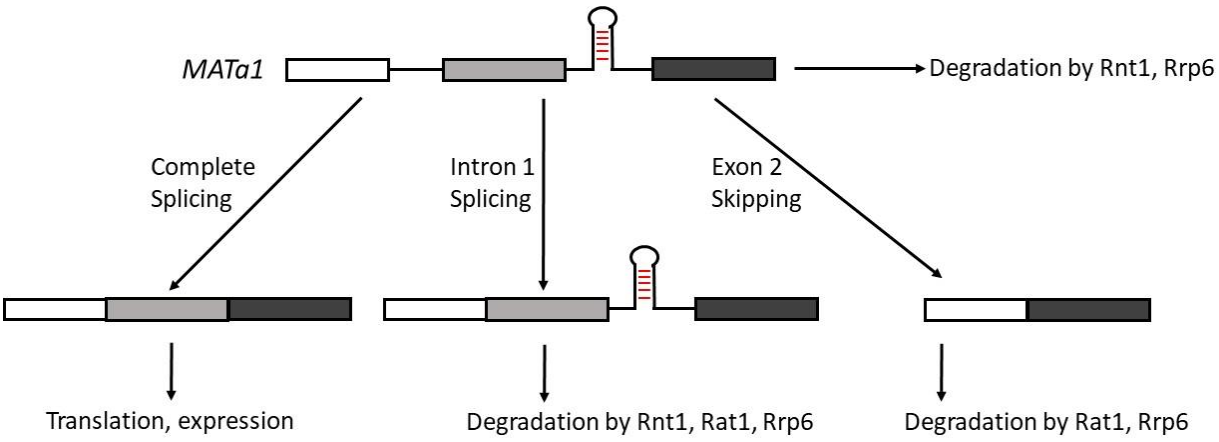


Figure 1.4. The two-intron transcript *MATa1* is regulated by multiple factors. The maturation or degradation of the *MATa1* transcript is dependent upon RNA splicing, the endonuclease Rnt1, and the nuclear exosome. The removal of both introns by the spliceosome yields a mature, translation-competent transcript. Unspliced or partially spliced species are subject to cleavage by Rnt1 and degradation by the nuclear exosome. The exon 2 skipped species is also a substrate for degradation by the nuclear exosome. Adapted from (Egecioglu et al., 2012).

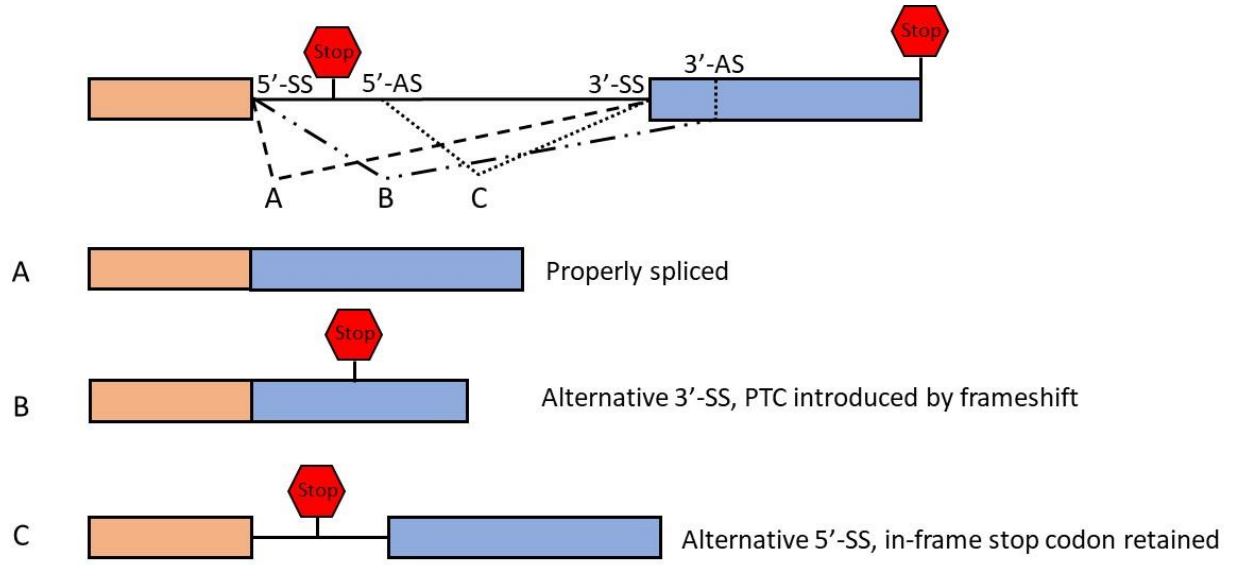
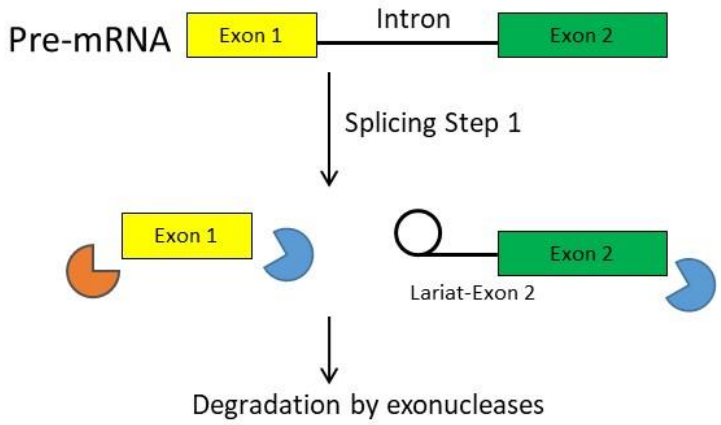


Figure 1.5. Alternative splicing may generate non-productive transcripts. A. Splicing at the annotated 5'- and 3'-SS generates a properly spliced mature transcript. **B.** Splicing at an alternative 3'-SS (3'-AS) downstream of the annotated 3'-SS may result in a translational frame shift that introduces a premature termination codon. **C.** Splicing at an alternative 5'-SS (5'AS) downstream of the annotated 5'-SS may lead to the inclusion of an intronic in-frame stop codon within the spliced transcript. Nonsense transcripts would then be subject to degradation by NMD.

A.



B.

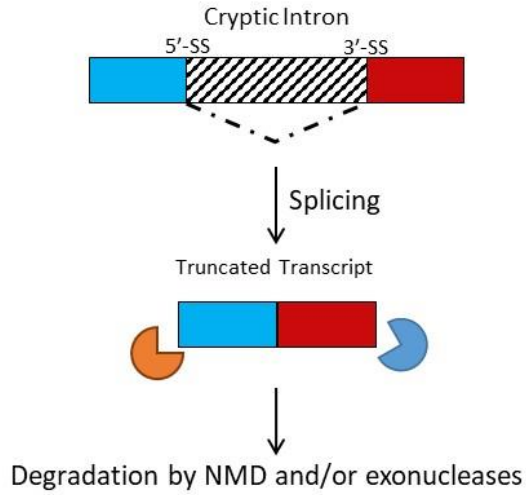


Figure 1.6. Spliceosome-mediated decay degrades partially spliced or cryptically spliced transcripts. A. The first catalytic step of splicing cleaves Exon 1 from the intron. If exon ligation is inefficient, the cleaved fragments are degraded by exonucleases. **B.** Certain RNA transcripts contain cryptic unannotated introns. The splicing of such an intron will produce a truncated transcript that may also be degraded by RNA surveillance pathways. If splicing yields a nonsense transcript then it may also be exported to cytoplasm, where it is subject to degradation by NMD.

References

Abovich, N., and Rosbash, M. (1997). Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. Cell *89*, 403–412.

Arenas, J.E., and Abelson, J.N. (1997). Prp43: An RNA helicase-like factor involved in spliceosome disassembly. PNAS *94*, 11798–11802.

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.

Ast, G. (2004). How did alternative splicing evolve? Nat. Rev. Genet. *5*, 773–782.

Awad, A.M., Venkataramanan, S., Nag, A., Galivanche, A.R., Bradley, M.C., Neves, L., Douglass, S., Clarke, C.F., and Johnson, T.L. (2017). Chromatin-remodeling SWI/SNF complex regulates coenzyme Q 6 synthesis and a metabolic shift to respiration in yeast. J. Biol. Chem. *292*, 14851–14866.

Barta, I., and Iggo, R. (1995). Autoregulation of expression of the yeast Dbp2p "DEAD-box" protein is mediated by sequences in the conserved DBP2 intron. EMBO J. *14*, 3800–3808.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. PNAS *74*, 3171–3175.

Bergkessel, M., Whitworth, G.B., and Guthrie, C. (2011). Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA *17*, 1461–1478.

Berglund, J.A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The Splicing Factor BBP Interacts Specifically with the Pre-mRNA Branchpoint Sequence UACUAAC. Cell *89*, 781–787.

Berglund, J.A., Rosbash, M., and Schultz, S.C. (2001). Crystal structure of a model branchpoint-U2 snRNA duplex containing bulged adenosines. RNA *7*, 682–691.

Black, D.L. (2003). Mechanisms of Alternative Pre-Messenger RNA Splicing. Annu. Rev. Biochem. *72*, 291–336.

Bonnet, A., Grosso, A.R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S.C., Janbon, G., Géli, V., de Almeida, S.F., and Palancade, B. (2017). Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. Mol. Cell *67*, 608–621.

Brys, A., and Schwer, B. (1996). Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3' splice site. RNA *2*, 707–717.

Burset, M., Seledtsov, I.A., and Solovyev, V. V (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res. *28*, 4364–4375.

Chan, S.P., and Cheng, S.C. (2005). The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. J. Biol. Chem. *280*, 31190–31199.

Chan, C.T.Y., Pang, Y.L.J., Deng, W., Babu, I.R., Dyavaiah, M., Begley, T.J., and Dedon, P.C. (2012). Reprogramming of tRNA modifications controls the oxidative stress response by codon-

biased translation of proteins. Nat. Commun. *3*, 937.

Chan, S.P., Kao, D.I., Tsai, W.Y., and Cheng, S.C. (2003). The Prp19p-Associated Complex in Spliceosome Activation. Science (80-.). *302*, 279–282.

Chao, J.A., and Williamson, J.R. (2004). Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. Structure *12*, 1165–1176.

Chapman, K.B., and Boeke, J.D. (1991). Isolation and characterization of the gene encoding yeast debranching enzyme. Cell *65*, 483–492.

Chen, H., and Cheng, S. (2012). Functional roles of protein splicing factors. Biosci. Rep. *32*, 345–359.

Chen, L., Muhlrad, D., Hauryliuk, V., Cheng, Z., Lim, M.K., Shyp, V., Parker, R., and Song, H. (2010). Structure of the Dom34–Hbs1 complex and implications for no-go decay. Nat. Struct. Mol. Biol. *17*, 1233–1240.

Chiu, Y.F., Liu, Y.C., Chiang, T.W., Yeh, T.C., Tseng, C.K., Wu, N.Y., and Cheng, S.C. (2009). Cwc25 Is a Novel Splicing Factor Required after Prp2 and Yju2 To Facilitate the First Catalytic Reaction. Mol Cell Biol *29*, 5671–5678.

Chow, L.T., Roberts, J.M., Lewis, J.B., and Broker, T.R. (1977a). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. Cell *11*, 819–836.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977b). An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell *12*, 1–8.

da Costa, P.J., Menezes, J., and Romão, L. (2017). The role of alternative splicing coupled to nonsense-mediated mRNA decay in human disease. Int. J. Biochem. Cell Biol. *91*, 186–175.

Crotti, L.B., and Horowitz, D.S. (2009). Exon sequences at the splice junctions affect splicing fidelity and alternative splicing. PNAS *106*, 18954–18959.

Dong, S., Li, C., Zenklusen, D., Singer, R.H., Jacobson, A., and He, F. (2007). YRA1 Autoregulation Requires Nuclear Export and Cytoplasmic Edc3p-Mediated Degradation of Its Pre-mRNA. Mol. Cell *25*, 559–573.

Egecioglu, D.E., Kawashima, T.R., and Chanfreau, G.F. (2012). Quality control of MATa1 splicing and exon skipping by nuclear RNA degradation. Nucleic Acids Res. *40*, 1787–1796.

Eng, F.J., and Warner, J.R. (1991). Structural basis for the regulation of splicing of a yeast messenger RNA. Cell *65*, 797–804.

Fourmann, J.B., Tauchert, M.J., Ficner, R., Fabrizio, P., and Lührmann, R. (2017). Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. Nucleic Acids Res. *45*, 4068–4080.

Frank, D., and Guthrie, C. (1992). An essential splicing factor, SLU7, mediates 3’ splice site choice in yeast. Genes Dev. *6*, 2112–2124.

Frank, D.N., Roiha, H., and Guthrie, C. (1994). Architecture of the U5 small nuclear RNA. Mol Cell Biol *14*, 2180–2190.

Gabunilas, J., and Chanfreau, G. (2016). Splicing-Mediated Autoregulation Modulates Rpl22p

Expression in Saccharomyces cerevisiae. PLoS Genet. *12*.

Gao, K., Masuda, A., Matsuura, T., and Ohno, K. (2008). Human branch point consensus is yUnAy. Nucleic Acids Res. *36*, 2257–2267.

Gowda, N.K.C., Kandasamy, G., Froehlich, M.S., Dohmen, R.J., and Andréasson, C. (2013). Hsp70 nucleotide exchange factor Fes1 is essential for ubiquitin-dependent degradation of misfolded cytosolic proteins. PNAS *110*, 5975–5980.

Gowda, N.K.C., Kaimal, J.M., Masser, A.E., Kang, W., Friedlander, M.R., and Andreasson, C. (2016). Cytosolic splice isoform of Hsp70 nucleotide exchange factor Fes1 is required for the degradation of misfolded proteins in yeast. Mol. Biol. Cell *27*, 1210–1219.

Hesselberth, J.R. (2013). Lives that introns lead after splicing. Wiley Interdiscip. Rev. RNA *4*, 677–691.

Hooks, K.B., Naseeb, S., Parker, S., Griffiths-Jones, S., and Delneri, D. (2016). Novel intronic RNA structures contribute to maintenance of phenotype in Saccharomyces cerevisiae. Genetics *203*, 1469–1481.

Horowitz, D.S. (2011). The splice is right: guarantors of fidelity in pre-mRNA splicing. RNA *17*, 551–554.

Horowitz, D.S. (2012). The mechanism of the second step of pre-mRNA splicing. Wiley Interdiscip. Rev. RNA *3*, 331–350.

Horowitz, D.S., and Abelson, J. (1993). Stages in the second reaction of pre-mRNA splicing: The final step is ATP independent. Genes Dev. *7*, 320–329.

Hossain, M.A., Claggett, J.M., Nguyen, T., and Johnson, T.L. (2009). The cap binding complex influences H2B ubiquitination by facilitating splicing of the SUS1 pre-mRNA. RNA *15*, 1515–1527.

Hossain, M.A., Rodriguez, C.M., and Johnson, T.L. (2011). Key features of the two-intron Saccharomyces cerevisiae gene SUS1 contribute to its alternative splicing. Nucleic Acids Res. *39*, 8612–8627.

Houseley, J., and Tollervey, D. (2009). The Many Pathways of RNA Degradation. Cell *136*, 763–776.

Howe, K.J., Kane, C.M., and Ares, M. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in *Saccharomyces cerevisiae*. RNA *9*, 993–1006.

Irimia, M., and Roy, S.W. (2014). Origin of spliceosomal introns and alternative splicing. Cold Spring Harb. Perspect. Biol. *6*.

James, S.A., Turner, W., and Schwer, B. (2002). How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. RNA *8*, 1068–1077.

Jiang, L., Whiteway, M., Ramos, C.W., Rodriguez-Medina, J.R., and Shen, S.H. (2002). The YHR076w gene encodes a type 2C protein phosphatase and represents the seventh PP2C gene in budding yeast [1]. FEBS Lett. *527*, 323–325.

Johnson, T.L., and Vilardell, J. (2012). Regulated pre-mRNA splicing: The ghostwriter of the eukaryotic genome. Biochim. Biophys. Acta *1819*, 538–545.

Juneau, K., Nislow, C., and Davis, R.W. (2009). Alternative splicing of PTC7 in Saccharomyces cerevisiae determines protein localization. Genetics *183*, 185–194.

Jurica, M.S., and Moore, M.J. (2003). Pre-mRNA splicing:awash in a sea of proteins. Cell *12*, 5–14.

Kawashima, T., Pellegrini, M., and Chanfreau, G.F. (2009). Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. RNA *15*, 2236–2247.

Kawashima, T., Douglass, S., Gabunilas, J., Pellegrini, M., and Chanfreau, G.F. (2014). Widespread Use of Non-productive Alternative Splice Sites in Saccharomyces cerevisiae. PLoS Genet. *10*.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. *13*, 700–712.

Köhler, A., Pascual-García, P., Llopis, A., Zapater, M., Posas, F., Hurt, E., and Rodríguez-Navarro, S. (2006). The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. Mol. Biol. Cell *17*, 4228–4236.

Konarska, M.M. (1998). Recognition of the 5 splice site by the spliceosome. Acta Biochim Pol *45*, 869–881.

Koodathingal, P., Novak, T., Piccirilli, J.A., and Staley, J.P. (2010). The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during Pre-mRNA splicing. Mol. Cell *39*, 385–395.

Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol. *14*, 153–165.

de Lima Morais, D.A., and Harrison, P.M. (2010). Large-scale evidence for conservation of NMD candidature across mammals. PLoS One *5*.

Lin, D., Yin, X., Wang, X., Zhou, P., and Guo, F.B. (2013). Re-annotation of protein-coding genes in the genome of saccharomyces cerevisiae based on support vector machines. PLoS One *8*.

Liu, Y.C., Chen, H.C., Wu, N.Y., and Cheng, S.C. (2007). A Novel Splicing Factor, Yju2, Is Associated with NTC and Acts after Prp2 in Promoting the First Catalytic Reaction of Pre-mRNA Splicing. Mol Cell Biol *27*, 5403–5413.

Lu, H., Zhu, Y.F., Xiong, J., Wang, R., and Jia, Z. (2015). Potential extra-ribosomal functions of ribosomal proteins in Saccharomyces cerevisiae. Microbiol. Res. *177*, 28–33.

Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat. Rev. Mol. Cell Biol. *16*, 665–677.

Ma, W.K., Cloutier, S.C., and Tran, E.J. (2013). The DEAD-box protein Dbp2 functions with the RNA-binding protein Yra1 to promote mRNP assembly. J. Mol. Biol. *425*, 3824–3838.

Marshall, A.N., Montealegre, M.C., Jiménez-López, C., Lorenz, M.C., and van Hoof, A. (2013). Alternative Splicing and Subfunctionalization Generates Functional Diversity in Fungal Proteomes. PLoS Genet. *9*.

Martin, A., Schneider, S., and Schwer, B. (2002). Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. J. Biol. Chem. *277*, 17743–17750.

Martín-Montalvo, A., González-Mariscal, I., Pomares-Viciana, T., Padilla-López, S., Ballesteros, M., Vazquez-Fonseca, L., Gandolfo, P., Brautigan, D.L., Navas, P., and Santos-Ocaña, C. (2013). The phosphatase Ptc7 induces coenzyme Q biosynthesis by activating the hydroxylase Coq7 in yeast. J. Biol. Chem. *288*, 28126–28137.

Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat. Rev. Mol. Cell Biol. *15*, 108–121.

Mathias, J.R., Hanlon, S.E., O'Flanagan, R.A., Sengupta, A.M., and Vershon, A.K. (2004). Repression of the yeast HO gene by the MATalpha2 and MATa1 homeodomain proteins. Nucleic Acids Res. *32*, 6469–6478.

Mayas, R.M., Maita, H., and Staley, J.P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat. Struct. Mol. Biol. *13*, 482–490.

McGlincy, N.J., and Smith, C.W.J. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem. Sci. *33*, 385–393.

Moore, M.J. (2002). Nuclear RNA Turnover. Cell *108*, 431–434.

Munding, E.M., Shiue, L., Katzman, S., Donohue, J., and Ares, M. (2013). Competition between Pre-mRNAs for the splicing machinery drives global regulation of splicing. Mol. Cell *51*, 338–348.

Nash, R., Weng, S., Hitz, B., Balakrishnan, R., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hirschman, J.E., et al. (2007). Expanded protein information at SGD: new pages and proteome browser. Nucleic Acids Res. *35*, D468–D471.

Newman, A.J. (1997). The role of U5 snRNP in pre-mRNA splicing. EMBO J. *16*, 5797–5800.

Newman, A.J., and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. Cell *68*, 743–754.

Nilsen, T.W. (1994). RNA-RNA interactions in the spliceosome: unravelling the ties that bind. Cell *78*, 1–4.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature *463*, 457–463.

O'Leary, M.N., Schreiber, K.H., Zhang, Y., Duc, A.C.E., Rao, S., Hale, J.S., Academia, E.C., Shah, S.R., Morton, J.F., Holstein, C.A., et al. (2013). The Ribosomal Protein Rpl22 Controls Ribosome Composition by Directly Repressing Expression of Its Own Paralog, Rpl22l1. PLoS Genet. *9*.

Ohrt, T., Prior, M., Dannenberg, J., Odenwalder, P., Dybkov, O., Rasche, N., Schmitzova, J., Gregor, I., Fabrizio, P., Enderlein, J., et al. (2012). Prp2-mediated protein rearrangements at the catalytic core of the spliceosome as revealed by dcFCCS. RNA *18*, 1244–1256.

Ohrt, T., Odenwälder, P., Dannenberg, J., Prior, M., Warkocki, Z., Schmitzová, J., Karaduman, R., Gregor, I., Enderlein, J., Fabrizio, P., et al. (2013). Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. RNA *19*, 902–915.

Parenteau, J., Durand, M., Véronneau, S., Lacombe, A.A., Morin, G., Guérin, V., Cecez, B.,

Gervais-Bird, J., Koh, C.S., Brunelle, D., et al. (2008). Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol. Biol. Cell *19*, 1932–1941.

Parenteau, J., Durand, M., Morin, G., Gagnon, J., Lucier, J.F., Wellinger, R.J., Chabot, B., and Elela, S.A. (2011). Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. Cell *147*, 320–331.

Parenteau, J., Lavoie, M., Catala, M., Malik-Ghulam, M., Gagnon, J., and Abou Elela, S. (2015). Preservation of Gene Duplication Increases the Regulatory Spectrum of Ribosomal Protein Genes and Enhances Growth under Stress. Cell Rep. *13*, 2516–2526.

Petibon, C., Parenteau, J., Catala, M., and Elela, S.A. (2016). Introns regulate the production of ribosomal proteins by modulating splicing of duplicated ribosomal protein genes. Nucleic Acids Res. *44*, 3878–3891.

Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Rapid, Transcript-Specific Changes in Splicing in Response to Environmental Stress. Mol. Cell *27*, 928–937.

Plocik, A.M., and Guthrie, C. (2012). Diverse forms of RPS9 splicing are part of an evolving autoregulatory circuit. PLoS Genet. *8*.

Raghunathan, P.L., and Guthrie, C. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Curr Biol *8*, 847–855.

Rain, J.C., and Legrain, P. (1997). In vivo commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein. EMBO J. *16*, 1759–1771.

Roecigno, R.F., Weiner, M., and Garcia-Blanco, M.A. (1993). A mutational analysis of the polypyrimidine tract of introns: Effects of sequence differences in pyrimidine tracts on splicing. J. Biol. Chem. *268*, 11222–11229.

Roy, K., and Chanfreau, G. (2014). Stress-induced nuclear RNA degradation pathways regulate yeast bromodomain factor 2 to promote cell survival. PLoS Genet. *10*.

Sayani, S., and Chanfreau, G.F. (2012). Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in Saccharomyces cerevisiae. RNA *18*, 1563–1572.

Sayani, S., Janis, M., Lee, C.Y., Toesca, I., and Chanfreau, G.F. (2008). Widespread Impact of Nonsense-Mediated mRNA Decay on the Yeast Intronome. Mol. Cell *31*, 360–370.

Schreiber, K., Csaba, G., Haslbeck, M., and Zimmer, R. (2015). Alternative Splicing in Next Generation Sequencing Data of Saccharomyces cerevisiae. PLoS One *10*.

Schwer, B. (2008). A Conformational Rearrangement in the Spliceosome Sets the Stage for Prp22-Dependent mRNA Release. Mol. Cell *30*, 743–754.

Schwer, B., and Meszaros, T. (2000). RNA helicase dynamics in pre-mRNA splicing. EMBO J. *19*, 6582–6591.

Scotti, M.M., and Swanson, M.S. (2015). RNA mis-splicing in disease. Nat. Rev. Genet. *17*, 19–32.

Shao, W., Kim, H.S., Cao, Y., Xu, Y.Z., and Query, C.C. (2012). A U1-U2 snRNP Interaction Network during Intron Definition. Mol Cell Biol *32*, 470–478.

Shaul, O. (2017). How introns enhance gene expression. Int. J. Biochem. Cell Biol. *91*, 145–155.

Smith, J.E., and Baker, K.E. (2015). Nonsense-mediated RNA decay-a switch and dial for regulating gene expression. BioEssays *37*, 612–623.

Spingola, M., Grate, L., Haussler, D., and Ares, M. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA *5*, 221–234.

Steffen, K.K., McCormick, M.A., Pham, K.M., Mackay, V.L., Delaney, J.R., Murakami, C.J., Kaeberlein, M., and Kennedy, B.K. (2012). Ribosome deficiency protects against ER stress in Saccharomyces cerevisiae. Genetics *191*, 107–118.

Stewart, M. (2010). Nuclear export of mRNA. Trends Biochem. Sci. *35*, 609–617.

Strauss, E.J., and Guthrie, C. (1994). PRP28, a “DEAD-box” protein, is required for the first step of mRNA splicing in vitro. Nucleic Acids Res. *22*, 3187–3193.

Tilgner, H., Knowles, D.G., Johnson, R., Davis, C.A., Chakrabortty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T.R., and Guigó, R. (2012). Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome Res. *22*, 1616–1625.

Tsai, R.T., Fu, R.H., Yeh, F.L., Tseng, C.K., Lin, Y.C., Huang, Y.H., and Cheng, S.C. (2005). Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes Dev. *19*, 2991–3003.

Tseng, C.K., Liu, H.L., and Cheng, S.C. (2011). DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. RNA *17*, 145–154.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34: Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3’ End of Aberrant mRNA. Mol. Cell *46*, 518–529.

Vilardell, J., and Warner, J.R. (1994). Regulation of splicing at an intermediate step in the formation of the spliceosome. Genes Dev. *8*, 211–220.

Volanakis, A., Passoni, M., Hector, R.D., Shah, S., Kilchert, C., Granneman, S., and Vasiljeva, L. (2013). Spliceosome-mediated decay (SMD) regulates expression of nonintronic genes in budding yeast. Genes Dev. *27*, 2025–2038.

Wang, G.S., and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nat. Rev. Genet. *8*, 749–761.

Warner, J.R., and McIntosh, K.B. (2009). How Common Are Extraribosomal Functions of Ribosomal Proteins? Mol. Cell *34*, 3–11.

Zhang, X., and Schwer, B. (1997). Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. Nucleic Acids Res. *25*, 2146–2152.

Chapter 2

Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*

OPEN ACCESS Freely available online

PLOS | GENETICS

Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*

Tadashi Kawashima[1¤], Stephen Douglass[2], Jason Gabunilas[1], Matteo Pellegrini[3], Guillaume F. Chanfreau[1]*

1 Department of Chemistry and Biochemistry and the Molecular Biology Institute, UCLA, Los Angeles, California, United States of America, **2** Bioinformatics Interdepartmental Program, UCLA, Los Angeles, California, United States of America, **3** Department of Molecular, Cellular and Developmental Biology, UCLA, Los Angeles, California, United States of America

Abstract

Saccharomyces cerevisiae has been used as a model system to investigate the mechanisms of pre-mRNA splicing but only a few examples of alternative splice site usage have been described in this organism. Using RNA-Seq analysis of nonsense-mediated mRNA decay (NMD) mutant strains, we show that many *S. cerevisiae* intron-containing genes exhibit usage of alternative splice sites, but many transcripts generated by splicing at these sites are non-functional because they introduce premature termination codons, leading to degradation by NMD. Analysis of splicing mutants combined with NMD inactivation revealed the role of specific splicing factors in governing the use of these alternative splice sites and identified novel functions for Prp17p in enhancing the use of branchpoint-proximal upstream 3′ splice sites and for Prp18p in suppressing the usage of a non-canonical AUG 3′-splice site in *GCR1*. The use of non-productive alternative splice sites can be increased in stress conditions in a promoter-dependent manner, contributing to the down-regulation of genes during stress. These results show that alternative splicing is frequent in *S. cerevisiae* but masked by RNA degradation and that the use of alternative splice sites in this organism is mostly aimed at controlling transcript levels rather than increasing proteome diversity.

Citation: Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF (2014) Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*. PLoS Genet 10(4): e1004249. doi:10.1371/journal.pgen.1004249

Editor: Hiten D. Madhani, University of California San Francisco, United States of America

Received October 8, 2013; **Accepted** February 4, 2014; **Published** April 10, 2014

Funding: Supported by grant GM061518 to GFC. TK was supported by USPHS National Research Service Award GM07104 and by a UCLA Graduate Division Dissertation Year Fellowship. JG was supported by a Ruth L. Kirschstein National Research Service Award GM007185. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: guillom@chem.ucla.edu

¤ Current address: Applied BioCode Inc., Santa Fe Springs, California, United States of America.

Introduction

Nonsense-mediated mRNA decay (NMD) is an RNA degradation system that degrades RNAs containing premature termination codons [1,2]. In mammalian cells and higher eukaryotes, NMD can be used to regulate gene expression, for instance by reducing the level of alternatively spliced isoforms containing premature termination codons [3,4,5,6,7,8]. This interplay between alternative splicing and NMD is involved in the autoregulation of SR proteins [3,4,5]. In addition to its function in regulating non-productively spliced isoforms, NMD is also used in a variety of eukaryotes to degrade unspliced pre-mRNAs that have escaped the splicing machinery [9,10,11,12,13]. Thus, NMD is widely involved in the proofreading of splicing efficiency and accuracy.

The yeast *Saccharomyces cerevisiae* has long been used as a model system to investigate the mechanisms of pre-mRNA splicing, as many components of the splicing machinery were identified through genetic screens in *S. cerevisiae* [14], and most splicing factors are highly conserved from yeast to mammalian cells [15]. Despite the presence of *c.a.* 330 intron-containing genes in *S. cerevisiae*, the prevalence of alternative splicing in this organism remains largely unexplored, as only a few examples of alternative splice site selection have been documented. The *SRC1* gene encodes an integral transmembrane protein, for which the use of an alternative 5′-splice site changes the number of passes through the membrane and ultimately the location of the C-terminal end of Src1p [16,17]. Alternative 3′-splice site selection has been shown to regulate expression of the *APE2* gene according to a temperature-dependent secondary structure in the transcript [18]. A few other alternative 3′-splice sites have been described, and the use of some of these sites produces transcripts that are degraded by NMD [19]. Recent work analyzing alternative splicing across fungal species has shown that *S. cerevisiae* has lost some of the alternative splicing events through gene duplication and sub-functionalization of the duplicated genes, which are otherwise produced by alternative splicing in other species [20].

Using RNA-Seq analysis of strains mutated for NMD factors, we identify here a large number of alternative splice sites in *S. cerevisiae*. However, we show that splicing at these sites is generally non-productive because it introduces premature termination codons (PTC), leading to degradation of the transcripts by NMD. Non-productive splicing can be increased during environmental stress to contribute to a global regulatory mechanism that down-regulates transcripts levels in response to environmental cues. These results show that non-productive splice sites are widely used in *S.cerevisiae*, but that transcripts spliced at these sites are

Author Summary

Accurate gene expression requires the transfer of gene information from DNA to RNA. When DNA is transcribed into RNA, part of the RNA needs to be removed (spliced) to generate a proper copy of the genetic information. This process needs to be very accurate to preserve the genetic information that will be transferred into proteins. Our study shows that in baker's yeast, the splicing process does not always produce the correctly spliced products, as RNA splicing events frequently utilize incorrect splice sites. However, these deficient RNA molecules are eliminated from cells by a quality control mechanism to preserve the integrity of the genetic information. However, incorrect splicing is not useless, as it can be used to regulate the quantity of RNA that is generated.

eliminated by RNA quality control mechanisms. Thus, while alternative splicing is frequently utilized in higher eukaryotes to generate proteome diversity, it is mainly used in *S.cerevisiae* as a means to regulate transcript levels.

Results

RNA-Seq reveals the accumulation of a large number of non-productive splice variants in NMD mutants

We previously showed that NMD degrades unspliced transcripts arising from a large fraction of intron-containing genes in *S. cerevisiae*, due to suboptimal splice sites [12,13], or upon splicing factor inactivation [21]. In addition, recent data showed that transcripts generated by the use of alternative 3′-splice sites can be degraded by NMD [19]. To gain further insights into the function of NMD in the proofreading of spliced isoforms, we performed RNA sequencing of mRNAs from wild-type and isogenic *upf1Δ*, *upf2Δ* and *upf3Δ* strains defective for NMD. To identify transcripts spliced at alternative splice sites, we performed gapped alignment analysis of the RNA sequences (Table S1) using BLAT [22]. This analysis revealed numerous occurrences of spliced transcripts arising from previously unknown splice sites, in both WT and the NMD mutants. We will refer to these new splicing events as alternative splicing events, even if these are found in wild-type cells, and to the annotated splicing events as the normal or canonical splicing events. Alternative splicing events were detected more frequently in RNA samples obtained from the NMD mutants (Fig. 1A and 1B; Table S2), consistent with the fact that most of these alternative splicing events result in the introduction of a PTC, either by inducing a translational frameshift or by inserting an intronic PTC-containing sequence (Table S2). After adjusting for sequencing depth, *upf1Δ*, *upf2Δ* and *upf3Δ* showed a 1.67, 1.72, and 1.90-fold enrichment in alternative splicing events and 1.59, 1.70, and 1.79-fold enrichment in PTC-generating alternative splicing events, respectively, versus wild-type (Table S2). NMD mutants showed an approximately 1.7-fold increase in unspliced mRNAs compared to the wild-type (Table S3) when considering reads that map to intronic and exon-intron regions, confirming our previous results from tiling arrays showing the involvement of NMD in eliminating unspliced transcripts genome-wide [12]. This enrichment for unspliced RNAs in NMD mutants is probably underestimated. Although there were 4-fold more reads that mapped only to intronic regions in NMD mutants compared to wild-type (Table S3), we observed an unanticipated high number of reads that mapped to exon-intron junctions in the wild-type strain (Table S3), which lowered the overall enrichment for unspliced RNAs in NMD mutants.

There was limited overlap in the alternative splicing events identified in the three *UPF* mutants (Fig. 1C), suggesting that the depth of our sequencing analysis was not sufficient to saturate identification of all alternative splicing events, particularly those occurring at low frequencies. The list of intron containing genes (ICG) for which we did not find the use of alternative splice sites is provided in Table S4, and the list of genes for which alternative splicing events were detected is shown in Table S5. 97 out of 304 intron containing genes analyzed did not exhibit alternative splicing (Table S4). Whether this reflects the absence of competing alternative sites or the lack of depth of our sequencing analysis remains to be determined.

To investigate if alternative splicing events are due to rare events or to splicing errors that occur randomly during transcript expression, we examined the abundance of ICG mRNAs that exhibited alternative splicing events and that of ICG mRNAs for which no alternative splicing events were detected (Fig. 1D). This analysis showed that some low abundance transcripts exhibited alternative splicing, while some high abundance transcripts did not (Fig. 1D). In addition, the median abundance of genes that showed alternative splicing was 117 RPKM, while the median abundance for genes with no alternative splicing events detected was 136 RPKM. Thus, even if the most highly expressed ICG (>2200 RPKM) all exhibited alternative splicing (Fig. 1D), genes with no alternative splicing were in general expressed at higher level than genes for which alternative splicing events were detected, showing no clear correlation between transcript abundance and the detection of alternative splicing events. We conclude that the detection of alternative splicing events in our RNA-Seq analysis is not an indirect consequence of the higher number of reads for highly-expressed transcripts.

The consensus sequences derived from the alternative splicing events identified in wild-type and all three mutants exhibited differences from the consensus sequences derived from the canonical (normal) splicing events (Fig. 1E). Alternative 5′-splice sites showed a relaxation of the conserved sequences, especially at positions 4 and 6 compared to the consensus obtained from the canonical splicing events. The 3′-splice sites used in alternative splicing events also showed a decrease in conservation of the polypyrimidine sequence preceding the conserved YAG, as well as a weaker conservation of the pyrimidine preceding the conserved AG dinucleotide (Fig. 1E). Thus, alternative splice sites identified by RNA sequencing showed a relaxed conservation, suggesting that these might correspond to lower efficiency splice sites, and possibly to regulated splicing events. Finally, we identified a number of alternative splicing events in either wild-type or NMD mutants that do not introduce a PTC and would potentially result in the production of proteins that differ from the SGD annotations. The list of these potential alternative proteins is presented in Table S6. However, we did not investigate these alternative protein forms further because most of the RNAs that would result in the production of these proteins were found in low abundance compared to those resulting in the production of the annotated proteins.

Strategy for validation of alternative splicing events

The previous RNA-Seq analysis revealed the potential wide-spread usage of alternative splice sites (SS). Figure 2 depicts specific mRNAs that were chosen for validation and further characterization. These transcripts were classified into three classes: those with 1) alternative 5′-SS; 2) alternative 3′-SS; and 3) a combination of both. Transcripts from class 1 included *RPL22B* as well as the previously reported *SRC1* [16]. Class 2 transcripts included genes encoding the RNA Polymerase III

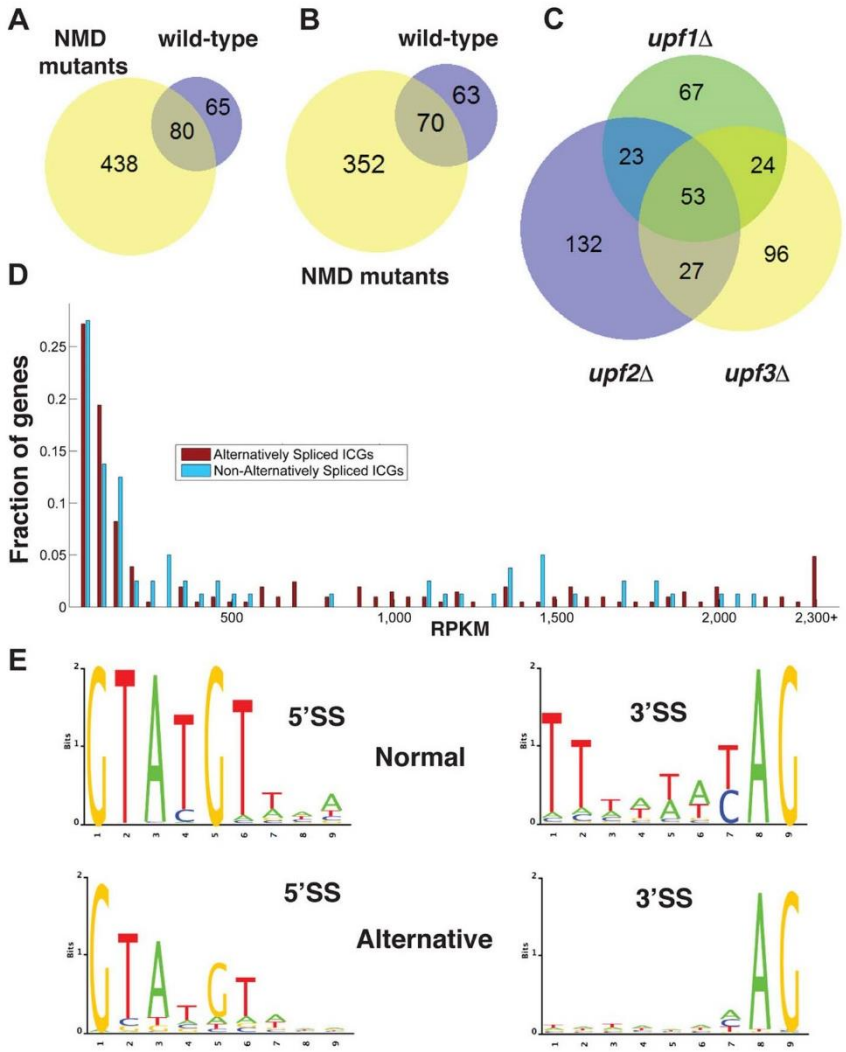


Figure 1. Bioinformatics analysis of alternative splice site usage in wild-type and NMD mutants. A. Venn diagram showing the overlap of alternative splice site usage between the wild-type and three NMD mutants pooled for all unique non-canonical splicing events (both PTC-generating and non-PTC-generating). B. Venn diagram showing the overlap of alternative splicing events between the wild-type and three NMD mutants pooled for all unique non-canonical splicing events resulting in a potential PTC. C. Venn diagram showing the overlap of alternative splicing events between the *upf1Δ*, *upf2Δ*, and *upf3Δ* strains for PTC-generating splicing events. D. Distributions of intron-containing gene transcripts showing alternative splicing events (red) or no alternative splicing events (blue) according to their overall abundance in RPKM. Transcripts for which the abundance was higher than 2,300 RPKM were grouped in the final bin. E. Sequence logo analysis of 5′- and 3′- splice sites for all normal and alternative splicing events detected by RNA-Seq in wild-type and NMD mutant strains.
doi:10.1371/journal.pgen.1004249.g001

transcription factor *TFC3* with a downstream alternative 3′-SS, and the adenosine deaminase *TAN1* with two alternative 3′-SS flanking the normal 3′-SS. For the third class, we examined genes encoding the glycosylphosphatidylinositol biosynthetic enzyme *GPI15* and the transcriptional regulator *GCR1*. *GPI15* exhibited the use of an alternative 5′-SS with the normal 3′-SS, as well as the normal 5′-SS with an alternative 3′-SS (Fig. 2). *GCR1* showed a more complex splicing pattern with multiple combinations of 5′ and 3′-SS (Fig. 2).

We analyzed alternative splicing events by RT-PCR using Cy3-end labeled primers, which allowed for relative comparison of the abundance of spliced and unspliced species, regardless of their size. Because we lacked an adequate size marker for Cy3 detection, the same RT-PCR analyses were initially performed with ^{32}P-end labeling with an appropriate ^{32}P-labelled ladder (data not shown) to confirm the sizes of all RT-PCR products and correlate the data back to gels obtained with Cy3-labeled primers. In addition to the wild-type and NMD-deficient *upf1Δ* strains, we analyzed the phenotypes of a number of *S. cerevisiae* splicing mutants. Knockout mutants of genes encoding Mud1p and Nam8p were chosen for their association with the U1 snRNP and role in 5′-SS selection [23,24,25,26]. The *HUB1* knockout was also included, as Hub1p was recently implicated in 5′-SS selection for *SRC1* [17]. Prp17p and Prp18p were selected for their involvement in the second step of splicing and potential effects on 3′-SS selection [27,28]. Finally, Isy1p was also included as a potential splicing fidelity factor [29]. The splicing profiles were analyzed for each of the genes

mentioned above in each of these mutant strains by fractionation of the RT-PCR products on polyacrylamide gels (Fig. 3). For the splicing mutants for which the splicing pattern differed from the wild-type, additional RT-PCR experiments were performed in triplicate from three independent cultures and quantitated, as shown in Supporting Figures S1, S2, S3, S4 and S6.

RT-PCR analysis confirms the involvement of Prp17p and Hub1p in *SRC1* alternative splicing

As a first step in validating our RT-PCR strategy, we focused on *SRC1*, which exhibits two possible 5′-SS (Fig. 2) and for which previous studies have demonstrated the roles of various splicing factors [16,17,30]. RT-PCR analysis of *SRC1* splice variants confirmed the use of these two alternative 5′-SS (Fig. 3). Wild-type samples showed a 60/40 ratio of *SRC1-S*/*SRC1-L* (Fig. S1), consistent with previous reports [16,17,30]. Samples from the *upf1Δ* mutant showed a pattern similar to wild-type (Fig. 3; Fig. S1), indicating that both variants are stable and not targeted by NMD. This result is consistent with our RNA-Seq analysis, which showed high sequence counts for both forms in all strains. Samples from the *nam8Δ* strain showed a slight increase in the level of unspliced transcripts (Fig. 3) due to reduced splicing efficiency [31]. The *prp17Δ* and *prp18Δ* mutants both showed a slight

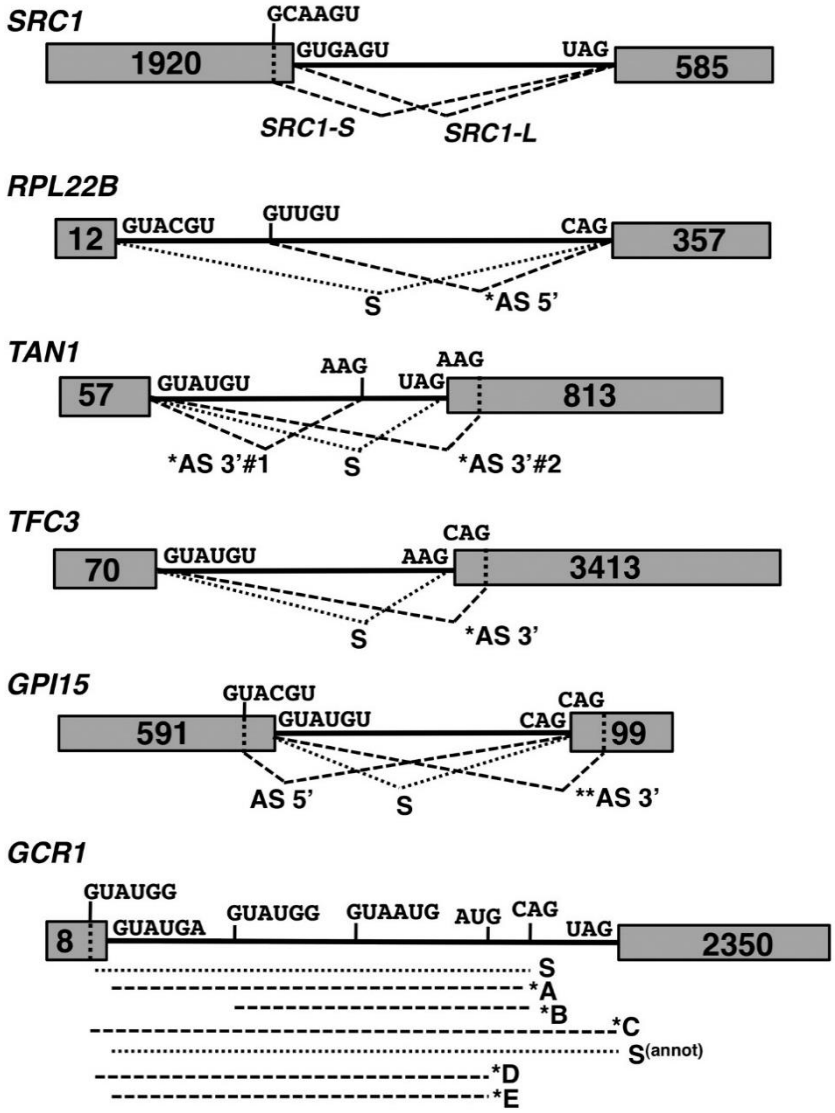


Figure 2. Spliced species produced from the *SRC1*, *RPL22B*, *TAN1*, *TFC3*, *GPI15* and *GCR1* genes. Species labeled with an asterisk are subject to NMD. Species labeled with two asterisks are predicted to be subject to NMD but were not observed to do so in subsequent experiments. The alternative 3′-SS of *SRC1* is located 4 nt upstream from the annotated 3′-SS. The alternative 3′-SS of *RPL22B* is located 64 nt downstream from the annotated 3′-SS. The alternative 3′-SS of *TAN1* are located 6 nt upstream and 7 nt downstream from the annotated 3′-SS. The alternative 3′-SS of *TFC3* is located 17 nt downstream from the annotated 3′-SS. The alternative 5′ and 3′-SS of *GPI15* are located 36 nt downstream and 14 nt upstream, respectively, from the annotated 5′ and 3′-SS. The alternative 3′-SS of *GCR1* are located 5 nt upstream (GUAUGG); 51 nt downstream (GUAUGG) and 627 nt downstream from the annotated 5′SS. The alternative 3′-SS of *GCR1* are located 40 nt upstream (AUG) and 17 nt downstream (CAG) from the annotated 3′-SS.

doi:10.1371/journal.pgen.1004249.g002

increase in the usage of *SRC1*-L 5′-splice site relative to the *SRC-S* 5′-splice site (1.4 and 1.3 fold, respectively), as suggested previously for the *prp17Δ* mutant at the protein level [17], and the *prp18Δ* mutant also exhibited an increase in unspliced precursors accumulation, consistent with previous results for other transcripts [21]. The *isy1Δ* mutant strain exhibited a clear accumulation of unspliced pre-mRNAs (Fig. 3), in agreement with the documented role of Isy1p in maintaining the proper conformation needed for the 1st step of splicing [29]. Hub1p inactivation resulted in a 3-fold reduction in the amount of *SRC1-S*, coinciding with an increase in *SRC1-L* (Fig. S1, Fig. 3), consistent with previous reports [17,30]. This reduction was also observed in the context of the *upf1Δ* mutant (Fig. S1, Fig. 3). Thus, the results described above confirmed the previously described effects of various splicing mutants on *SRC1* splicing patterns and showed that our RT-PCR strategy is effective in analyzing the impact of specific splicing factors on splice site usage..

Efficient use of the non-productive 5′-splice site of *RPL22B* is strongly dependent on the U1 snRNP components Nam8p and Mud1p

RPL22B showed the presence of an alternative 5′-SS in the intronic sequence, which unlike *SRC1*, yields a PTC-containing transcript potentially targeted to NMD (Fig. 2). This alternatively spliced transcript was almost 10-fold more abundant in the *upf1Δ* mutant (Fig. 3; Fig. S2), further suggesting that it is targeted by NMD. We also detected a large accumulation of unspliced species

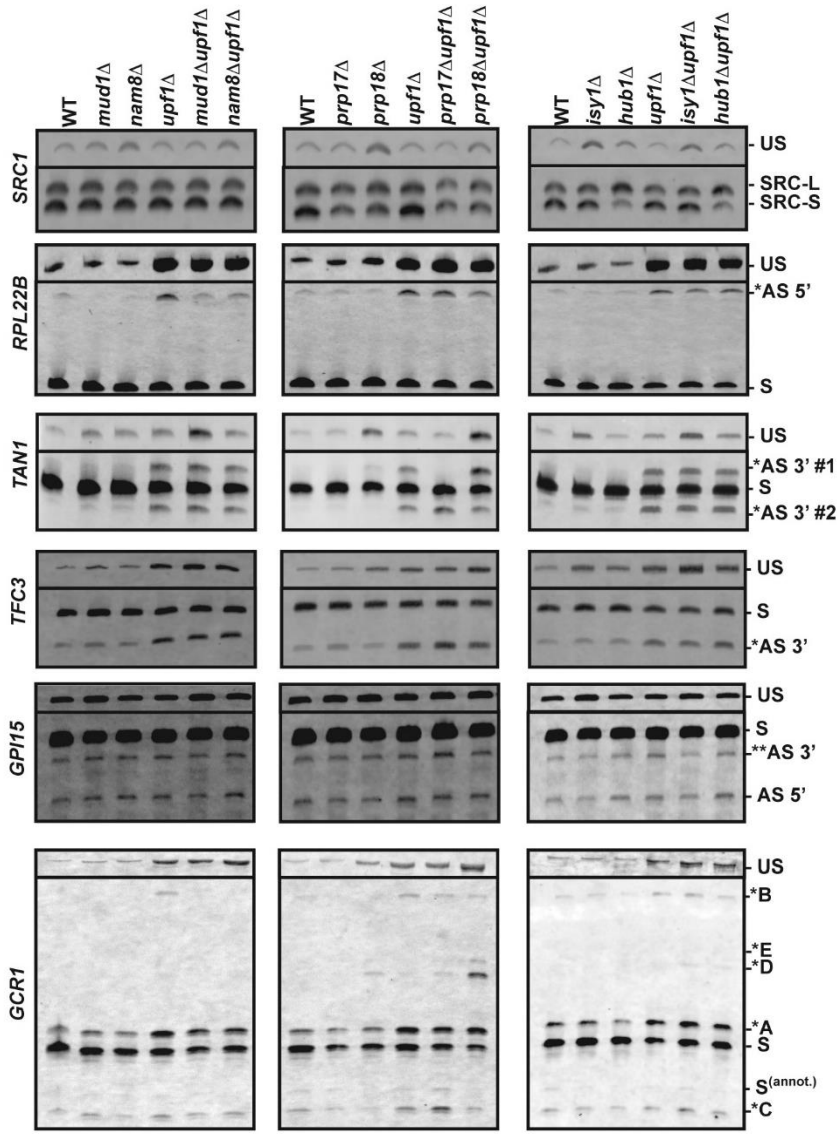


Figure 3. RT-PCR analysis of alternatively spliced products for *SRC1*, *RPL22B*, *TAN1*, *TFC3*, *GPI15* and *GCR1* in wild-type, NMD and various splicing mutants. The unspliced (US) species is also shown on top. The middle portions of the gel where no species were visible have been removed. In all cases, RT-PCR was performed with a Cy3-labeled primer. The labeling of the different alternatively spliced forms is according to the nomenclature shown in Figure 2.
doi:10.1371/journal.pgen.1004249.g003

in the *upf1Δ* mutant, indicating inefficient recognition of this splicing substrate. This may be the result of both the normal (GUACGU) and alternative (GUUUGU) 5′-SS having non-consensus sequences (see below). Interestingly, the abundance of the alternatively spliced product was found to decrease by two to three folds when Nam8p or Mud1p were inactivated in the context of the *upf1Δ* deletion (Fig. 3; Fig. S2). The deletion of either one of these two factors might hinder the ability of the U1 snRNP to bind to the alternative suboptimal 5′ GUUUGU splice site of *RPL22B*, resulting in decreased usage. This is consistent with the known roles of Mud1p and Nam8p in the first step of splicing [25], and suggest their direct involvement in modulating 5′-SS selection of *RPL22B*. By contrast, no major changes were observed in the *prp17Δ*, *prp18Δ*, *isy1Δ*, *hub1Δ* mutants, either alone or in combination with the *upf1Δ* deletion (Fig. 3), showing the specificity of the effects detected with Nam8p and Mud1p. Thus, *RPL22B* exhibits two competing suboptimal 5′-SS, one of which is highly sensitive to perturbations in the U1 snRNP. The functional significance of the alternative 5′-SS of *RPL22B* in regulating transcript levels is investigated further below.

A novel role for Prp17p in promoting the use of branchpoint proximal alternative 3′-splice sites

Gapped sequence alignment showed that *TFC3* exhibits an alternative CAG 3′-SS 17 nt downstream of the annotated AAG (Fig. 2). This product can be detected in samples from the wild-type and splicing mutants, but is 4.5-fold more abundant in the context of the *upf1Δ* deletion, showing that a large fraction of this product is degraded by NMD (Fig. 3; Fig. S3). This non-productive isoform amounts to 27% of all spliced products (Fig. S3), showing that a significant fraction of splicing generates NMD-targeted, non-productive transcripts. We observed a slight accumulation (1.7 fold) of the downstream alternative 3′-splice product in the *prp17Δ* mutant (Fig. 3; Fig. S3), showing that this second step splicing factor contributes to reducing the use of this alternative 3′-SS. As expected, inactivation of the first step splicing factors Mud1p or Nam8p had no effect on the pattern of 3′-SS selected (Fig. 3).

TAN1 exhibits a more complex alternative 3′-SS pattern, where a canonical UAG 3′-SS is flanked by two alternative 3′ AAG sequences (Fig. 2). The use of either of these sites would generate PTC-containing transcripts. The upstream AAG (AS 3′ #1) is only 6 nt away from the canonical 3′-SS. The retention of 6 nt of intronic sequence would maintain the proper reading frame but would result in a PTC because the UAG sequence of the normal 3′-splice site corresponds to an in-frame stop codon [32]. The downstream AAG (AS 3′ #2) is 7 nt downstream of the normal 3′-SS, resulting in a frameshift-induced PTC. RT-PCR analysis of the wild-type and *upf1Δ* strains confirmed the RNA-Seq data by showing that these two alternative splice products are detected at extremely low levels, unless NMD is inhibited (Fig. 3; Fig. S4). In samples from the *upf1Δ* strain, the two alternatively spliced products accumulate to similar amounts, and both species are detected at lower levels than the normal spliced product (20% of all spliced products; Fig. S4), possibly because these two suboptimal AAG sites do not compete well with the consensus canonical UAG site. Strikingly, the usage of these alternative 3′ splice sites was dramatically altered when Prp17p or Prp18p were inactivated. Inactivating Prp17p resulted in an increase in the use of the downstream alternative 3′-SS (AS 3′#2), while the upstream alternative 3′-SS (AS 3′#1) was no longer used (Fig. 3; Fig. S4), showing a role of Prp17p in enhancing the use of upstream, branchpoint proximal 3′-SS. By contrast, Prp18p inactivation resulted in increased usage of the alternative 3′-SS most proximal to the branch point sequence (AS 3′ #1; Fig. 3; Fig. S4). This product is barely detectable in the wild-type strain but can be observed in the *prp18Δ* strain (Fig. 3), and inactivation of Prp18p in the context of the *upf1Δ* deletion resulted in a 3-fold increase in the abundance of this species (Fig. S4). The effect of Prp18p on this 3′-SS might be due to the identity of the sequences immediately following the 3′-SS, which have been shown to influence 3′-SS selection in the absence of a functional Prp18p [33]. Isy1p inactivation resulted in an increase of unspliced species in a similar fashion to *SRC1* discussed above; however there was no effect of Isy1p, Hub1p, Mud1p and Nam8p on alternative 3′-SS selection of *TAN1* (Fig. 3), showing the specificity of the effects observed with Prp17p and Prp18p. Finally, unspliced *TAN1* transcripts were generally not affected by NMD, except in the context of a *mud1Δ* mutant strain (Fig. 3). This observation is consistent with a recent report showing that *TAN1* unspliced transcripts are retained in the nucleus by the RES complex, and are subject to NMD only when the RES complex is inactivated [34]. Overall, analysis of *TFC3* and *TAN1* alternative 3′-SS patterns show that Prp17p and Prp18p have antagonistic roles in the selection of upstream and downstream 3′-SS *of TAN1*, and highlight the importance of Prp17p in enhancing the use of 3′-SS located closer to the branchpoint.

Alternative splicing patterns of *GPI15* and *GCR1* reveal the production of alternative non-functional protein products and the use of a non-canonical AUG 3′-splice site repressed by Prp18p

GPI15 in an interesting case where the two alternatively spliced products identified by our RNA-Seq analysis are not targeted by NMD. The use of an alternative GUACGU 5′-splice site results in the deletion of 30 nucleotides from the 3′ end of exon 1 (Fig. 2), which maintains the open-reading frame but generates a truncated protein. However, the protein product resulting from translation of this alternatively spliced product is likely to be non-functional, as this truncation removes a stretch of 10 amino acids at positions 187–197 in the most highly conserved region of this protein [35]. This transcript can be detected in samples from the wild-type and the splicing factor mutants, and does not vary in intensity in the context of *upf1Δ*, indicating that it is not targeted by NMD (AS 5′, Fig. 3). In contrast, the alternatively spliced transcript generated by use of a downstream CAG 3′-SS results in a PTC. However, this PTC-containing transcript would exhibit a short 85 nt 3′-UTR, which might render it insensitive to NMD as suggested by the *faux 3′ UTR* model [36,37]. Indeed, the abundance of this transcript was not increased in the *upf1Δ* mutant (Fig. 3). In addition, this transcript is expected to yield a non-functional protein due to C-terminal truncation and deletion of amino acids within the most conserved region of the protein [35]. Analysis of the pattern of selection of these two alternatively spliced transcripts in the various splicing mutants did not reveal any major effect of these mutants (Fig. 3) in contrast to the effects described above for *RPL22B*, *TAN1* or *TFC3*. However, there was a slight increase in the use of the downstream alternative 3′-SS in the *prp17Δupf1Δ* strain, consistent with the role of Prp17p in favoring the upstream 3′-SS, as described above for *TAN1* and *TFC3*.

GCR1 showed the most complex splicing pattern of all transcripts analyzed. Gapped alignments identified an intronic GUAUGG alternative 5′-SS as well as an upstream CAG alternative 3′-SS (Fig. 2). In addition to these alternative splice sites identified by RNA-Seq, RT-PCR revealed the use of an additional GUAUGG alternative 5′-SS staggered 5-nt upstream of the normal 5′-SS and of a non-canonical AUG alternative 3′-SS

23 nt upstream of the other alternative 3′-SS (Fig. 2). The use of all of these sites was confirmed by RT-PCR, cloning and Sanger sequencing (see below and Fig. S5). The fact that some alternative splice sites escaped identification by mRNA sequencing indicates that a greater depth of coverage has the potential to identify even more alternative splice sites.

Based on *GCR1* annotation, the canonical spliced mRNA would use the GUAUGA 5′-SS along with the most downstream UAG 3′-SS (Fig. 2). This product (labeled as $S^{(annot.)}$ in Fig. 2 and 3), however was detected at very low levels (Fig. 3). The major spliced product observed resulted from the use of the most upstream GUAUGG 5′-SS and of an upstream CAG 3′-SS (labeled "S" in Fig. 2 and 3). This splicing event does not introduce a PTC and results in a protein that is very similar to the translation product of the annotated spliced transcript $S^{(annot.)}$. The annotated amino acid sequence of *GCR1* from position 2 to 4 is VCT. In the major spliced product S, this sequence is replaced by QTSVDST. Thus, most of the protein is identical, except for a few N-terminal amino acids which are not expected to affect Gcr1p function, as all *GCR1* mutations with phenotypic effects have been mapped to a region downstream of this short sequence stretch [38,39,40]. Based on the relative abundances of S and $S^{(annot.)}$, it is clear that S, and not $S^{(annot.)}$ is the main spliced product for the *GCR1* gene.

In addition to this major spliced product, we also detected a series of alternatively spliced products degraded by NMD (as denoted by asterisks in Fig. 2 and 3). Splicing from the annotated GUAUGA 5′ splice site combined with the upstream CAG 3′ splice site resulted in a PTC-containing transcript labeled as *A in Figs. 2 & 3. This transcript is degraded by NMD, as higher amounts are observed in all the strains containing a *upf1Δ* deletion, and it is the most abundant of all *GCR1* alternatively spliced products subject to NMD (Fig. 3; Fig. S6). Another product is generated from combining the upstream GUAUGG 5′-SS with the most downstream UAG 3′-SS (*C in Fig. 2). This splicing event results in a PTC, as it introduces a translational frameshift, which is not detected until the 43rd amino acid is translated. The corresponding transcript accumulates at low abundance in all samples and appears to be targeted by NMD, as its abundance increases slightly in all *upf1Δ* strains. In addition, the use of this most downstream 3′-SS increases almost 4-fold in the *prp17Δupf1Δ* strain when compared to the *upf1Δ* control (Fig. 3; Fig. S6). Because the 3′-SS used to generate this transcript corresponds to the most downstream one, this observation provides another example of the importance of Prp17p in favoring the selection of upstream 3′-SS, as shown above for *TFC3*, *TAN1* and to a lesser extent *GPI15*.

Another PTC-containing transcript that is degraded by NMD results from splicing of the downstream intronic GUAUGG 5′-SS with the CAG 3′-SS, (labeled *B in Fig. 2 and 3). This product is faint, but detectable in all cases of NMD deactivation, except in combination with *nam8Δ* or *mud1Δ*, most likely because this 5′-SS has a higher sensitivity to U1 snRNP perturbations, as described above for *RPL22B*. Analysis of other mutants did not reveal any major influence on the pattern of 5′- or 3′-SS selection. Like *SRC1*, *GCR1* exhibits two staggered 5′ splice sites. However, unlike for *SRC1*, Hub1p has no influence on their selection (Fig. 3).

A final set of NMD targets are produced by the use of the two most upstream 5′-SS with a highly unusual alternative AUG 3′-SS in the intronic sequence (labeled *D and *E in Figs. 2 &3). Interestingly, these products were only detected in the absence of Prp18p, suggesting that this factor is essential in preventing the use of this non-canonical 3′-SS. The use of this highly unusual AUG 3′ splice site was unambiguously confirmed through sequencing and RT-PCR analysis of RNAs derived from *prp18Δupf1Δ* samples (Fig. S5). The ATPase Prp22p has been implicated in the fidelity of 3′-SS selection [41]. Because Prp18p functions upstream from Prp22p during the late stages of splicing [42], it is possible that the absence of Prp18p might indirectly hinder the function of Prp22p in proofreading 3′-SS selection, and that the use of this unusual 3′-SS might be the consequence of a reduced Prp22p function in the absence of Prp18p. To test this hypothesis, we analyzed *GCR1* splicing in a *prp22-1* mutant. RT-PCR analysis showed that the spliced product generated from the use of the AUG 3′-SS did not accumulate in a *prp22-1* splicing mutant (Fig. S7). Thus, the accumulation of species resulting from the use of this unusual 3′-SS in the *prp18Δupf1Δ* samples is not an indirect consequence of hindered Prp22p function. The discovery of the splicing at this unusual 3′-SS sequence reveals the importance of Prp18p in ensuring proper 3′-SS selection for *GCR1* and in repressing the use of non-canonical 3′-SS sequences.

Alternatively spliced species of *RPL22B* and *GCR1* increase during stress conditions

The previous results validated our prediction that transcripts generated from the use of alternative non-productive splice sites are degraded by NMD and revealed the role of specific splicing factors in governing the choice between alternative sites. Strikingly, the sequence of some of these non-productive splice sites was found to be conserved across closely related yeast species (Fig. S8, *RPL22B* and Fig. S9, *TAN1*). Because the level of sequence conservation in intronic sequences is usually very low (Fig. S8, S9), these peaks in sequence conservations for intronic alternative splice sites might reflect their functional importance. We hypothesized that the use of some of these alternative splice sites which lead to degradation by NMD might be favored under certain conditions to down-regulate gene expression. To test this hypothesis, we monitored changes in the splicing patterns of *RPL22B*, *TAN1*, and *TFC3* under stress conditions such as amino acid starvation, heat shock, LiCl-mediated hyperosmotic stress, and rapamycin treatment, as these have been reported to elicit diverse responses in the expression of intron containing genes [43,44]. In addition, various stresses cause down-regulation in ribosomal protein gene expression (many of which contain introns), presumably to relieve the cell of massive energy requirements of ribosome biogenesis and focus those resources into regulations that are the most appropriate in response to the current stress condition [45,46,47]. After 10 minutes of amino acid depletion, *RPL22B* showed an increase in unspliced species as well as well as a 4.5-fold increase in the level of the alternatively spliced product when compared to the SDC or YPD media controls (Fig. 4A; Fig. S10). In the *upf1Δ* strain shifted to amino acid starvation conditions, the levels of the alternatively spliced product increased compared to the wild-type strain grown in the same conditions, as would be expected when NMD transcripts are no longer degraded (Fig. 4A lanes 2 and 4). The fact that the level of the alternatively spliced transcript is 2.5-fold higher in the *upf1Δ* sample than in the wild-type sample under amino acid starvation conditions (Fig. S10) argues that the increase in the abundance of these species in the wild-type strain in these conditions is not due to NMD inhibition in these conditions, but that a change in splice site selection occurs that favors the use of the alternative splice site. Significantly, amino acid starvation did not change the levels of the alternatively spliced species of *TAN1* and *TFC3* that are normally subject to NMD (Fig. S11). This observation provides further evidence that the increase in the amount of alternatively spliced *RPL22B* transcript observed during amino acid starvation is due to a switch in splice site selection and not to an inhibition in NMD, since the level of alternatively spliced species of *TAN1* and *TFC3* that are normally degraded by NMD is unaffected in the same conditions.

We next investigated the effect of a 20 minute heat shock at 42°C on splicing patterns. Under these conditions and in the wild-type strain, *RPL22B* showed an increase in unspliced as well as a decrease in the relative amount of the normal spliced product (Fig. 4A lane 5 vs. 7). More importantly, the NMD defective strain *upf1Δ* showed an even larger increase in unspliced pre-mRNAs, as well a large accumulation of the alternatively spliced product that coincides with a decreased amount of canonical spliced product (Fig. 4A lane 6 vs. 8). In these conditions, the alternatively spliced product now corresponds to more than half of all spliced species (Fig. S10). Under heat shock, this alternatively spliced product is 4-fold more abundant in the *upf1Δ* strain than in the wild-type strain. These higher levels upon NMD inactivation show that the increased accumulation of these species under heat shock is not due to a decrease in NMD efficiency. Rather, this result shows that the use of the alternative splice site is being favored in heat shock conditions. By contrast, *TFC3* and *TAN1* exhibited an accumulation of unspliced species, but decreased levels of both the canonical and alternatively spliced species (Fig. S11), consistent with a general inhibition of pre-mRNA splicing under heat shock [48,49]. Thus the pattern of alternatively spliced species of *TFC3* and *TAN1* that are subject to NMD is very different from that of *RPL22B*, further proving that the accumulation of the alternatively spliced *RPL22B* transcript under heat shock conditions described above is not due to a general stabilization of spliced forms degraded by NMD.

Like heat shock, rapamycin treatment was shown to result in an inhibition of ribosomal proteins mRNA splicing based on microarray experiments [43]. Within 20 minutes of rapamycin treatment, *RPL22B* indeed showed trends similar to those observed in heat shock, but to a lesser degree, with an increase of unspliced species and of alternatively spliced *RPL22B* species (Fig. 4A), but no effect on the alternatively spliced *TAN1* and *TFC3* transcripts (Fig. S11). Hyperosmotic shock (300 mM LiCl exposure for 10 min) only resulted in minimal effects; there were no changes observed for *TFC3* and *TAN1* targets under these stress conditions (Fig. S11), and *RPL22B* showed only a slight increase in unspliced but the levels of spliced transcripts remained similar. Thus, *RPL22B* exhibits regulated use of its alternative 5′-splice site, mostly under amino acid starvation and heat shock conditions, while other transcripts such as *TFC3* and *TAN1* did not exhibit any change in their alternative splicing profiles.

Because *GCR1* exhibited a very complex splicing pattern, especially in the absence of Prp18p, and because heat shock conditions resulted in the most dramatic changes in splicing for *RPL22B*, we next investigated the effect of heat shock on *GCR1* splicing in the wild-type, *upf1Δ*, *prp18Δ* and *prp18Δupf1Δ* mutants (Fig. 4B). Under heat-shock, we detected a general inhibition of splicing, consistent with the data described above. However, we also observed an increase of the abundance of the A* form relative to the normal spliced product S, indicative of a switch from the normal GUAUGG site to the GUAUGA site. The absence of Prp18p resulted in an increase of the use of the non-canonical AUG site (*D species) in heat shock conditions, and this product now constituted one third of all spliced species. Thus we conclude that *GCR1*, like *RPL22B*, exhibits a switch in splice site selection during heat shock, and that Prp18p limits splicing at this non-canonical AUG site under stress conditions.

The alternative, suboptimal 5′-splice site of *RPL22B* contributes to the global down-regulation of *RPL22B* in stress conditions

To further analyze the importance of the alternative 5′-splice site of *RPL22B* on its splicing patterns and expression during normal and stress conditions, we investigated the effect of mutations of this alternative 5′-SS. The suboptimal GUUUGU alternative 5′-splice site was either deleted or mutated to the consensus GUAUGU sequence at the endogenous chromosomal locus (CS, consensus mutation and **Δ**, deletion, Fig. 5A). Changing the alternative 5′-SS to the consensus GUAUGU sequence resulted in detectable amounts of alternatively spliced products at 25°C, even in a functional NMD background (Fig. 5B), suggesting that the suboptimal GUUUGU sequence contributes to the low usage of this alternative site in normal conditions. Inactivation of Upf1p in this context showed that 70% of all spliced species were now being produced by splicing from the alternative consensus site (Fig. 5B; lane 5; Fig. S12 for quantitation), and that splicing efficiency was improved, as shown by a decrease in unspliced species. By contrast, deleting the alternative splice site resulted in higher amount of unspliced transcripts, especially in the *upf1Δ* background. Thus, deleting the alternative 5′-splice site of *RPL22B* is not sufficient to improve splicing at the normal splice site, possibly because of the suboptimal sequence of the normal *RPL22B* 5′-splice site. In addition to RT-PCR, the same strains were analyzed by northern blot (Fig. 5B, bottom panel), which yielded results similar to those obtained by RT-PCR. These results show that increasing the strength of the alternative 5′-SS of *RPL22B* is sufficient to enhance the overall splicing efficiency of this transcript, while deleting this site results in an overall increase of unspliced RNAs. Under heat shock and NMD inactivation, this effect was even more prominent, as mutation of the alternative splice site to the consensus resulted in the alternatively spliced product being the major spliced species (Fig. 5B, lane 11). Thus, under heat shock conditions, *RPL22B* transcripts bearing the consensus alternative splice site mutation are now spliced almost exclusively at this site. Analysis of the mutant with a deletion of the alternative 5′-SS under heat shock conditions showed that the use of the normal 5′-SS is not increased at elevated temperatures when the competing alternative 5′-SS has been eliminated (Fig. 5A and 5B lane 12). This mutant shows a larger accumulation of unspliced *RPL22B* transcript, hinting that the normal process of spliceosome assembly is perturbed on this transcript during heat shock, possibly due to the suboptimal 5′-SS. To obtain a more quantitative assessment of transcript levels, rather than just assessing the ratio between the different spliced forms, we analyzed the same samples by northern blot. This analysis showed that cells treated in heat shock conditions resulted in much weaker signal than in the samples obtained from cells grown at 25°C, consistent with a general down-regulation of ribosomal protein genes under stress [43,44,45,46,47]. Upon NMD inactivation, we observed a rescue of transcript levels, which mostly corresponded to unspliced RNAs and to some alternatively spliced transcripts (Fig. 5B). However, changing the alternative 5′-SS to a consensus sequence in the context of NMD inactivation was sufficient to recover a large amount of spliced transcripts (Fig. 5B, lane 11, lower panel). To investigate if this effect was specific to heat shock or is also observed during other stresses, we analyzed the expression of wild-type and mutated forms of *RPL22B* during amino acid starvation (Fig. 5C). The results observed during amino acid starvation were similar to those described during heat shock, with a large increase in the level of spliced transcripts upon changing the alternative 5′-splice site to the consensus sequence. We also observed an increase in the use of the alternative 5′-SS under amino acid starvation (Fig. S13). Interestingly, shifting the Upf1p-inactivated strain with the alternative 5′-splice site consensus sequence from SDC to amino acid starvation conditions resulted in only a minor increase of the use of the alternative 5′-SS, possibly because the level of transcripts spliced at that site is already very high in the *upf1Δ*

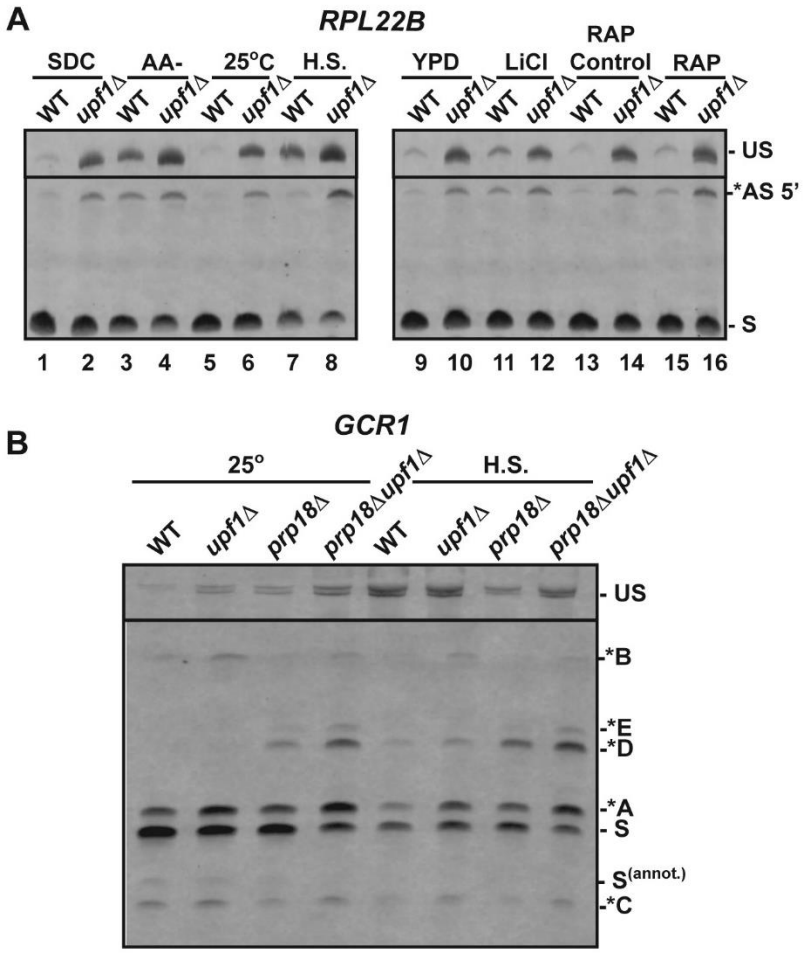


Figure 4. RT-PCR analysis of alternatively spliced products under stress conditions. A. Analysis of *RPL22B* in various stress conditions. Shown are the RT-PCR products obtained from the wild-type or *upf1Δ* mutant strain after growth in the following conditions: SDC, synthetic define complete medium at 30°C; -AA, 10 minutes in SDC medium at 30°C lacking amino acid (-AA); 25°C, log phase at 25°C in YPD; H.S, 20 minutes at 42°C in YPD; YPD: log phase at 30°C in YPD; LiCl, incubation with 300 mM LiCl in YPD at 30°C for 10 minutes; RAP control, see Materials and Methods; RAP, treatment with Rapamycin for 20 minutes. B. RT-PCR analysis of *GCR1* alternative splicing in heat-shock conditions. Labeling of the different species is similar to that of Figures 2 and 3.
doi:10.1371/journal.pgen.1004249.g004

strain in normal conditions (75%; Fig. S13). In conclusion, these results show that low splicing efficiency due to the suboptimal normal and alternative 5′ splice sites of *RPL22B*, combined with NMD degradation of the unspliced and alternatively spliced forms contribute to the general decrease in *RPL22B* levels as a means to rapidly halt production of this ribosomal protein under various stress conditions.

Usage of the alternative 5′-splice site of *RPL22B* is influenced by promoter identity

Ribosomal protein genes are known to be transcriptionally regulated in stress conditions. To investigate the use of *RPL22B* 5′-SS selection independently from transcriptional inhibition under heat shock, we replaced the natural *RPL22B* promoter with a galactose-inducible promoter. The wild-type and *upf1Δ* strains containing the natural *RPL22B* promoter showed no detectable difference in *RPL22B* splicing patterns or expression when grown in galactose containing medium (YPGal) compared to glucose-containing medium (YPD) at 25°C (Fig. 6 lanes 1–4), either by RT-PCR (top panel) or northern blot (bottom panel). Strikingly, replacement of the normal *RPL22B* promoter by the *GAL* promoter resulted in an increase in overall *RPL22B* transcript levels, but also in a decrease in the use of the alternative 5′-SS (Fig. 6). The fact that the usage of the alternative splice site of *RPL22B* is reduced in this strain while transcript levels are higher overall argues against the hypothesis that alternative splice site usage is the result of splicing errors occurring at low frequencies, as if this were the case, one would expect higher levels of alternatively spliced *RPL22B* transcripts upon its overexpression in the strain in which the natural *RPL22B* promoter was swapped for the *GAL* promoter. Under heat shock conditions, the use of the alternative splice site was reduced 8.1 fold in the *upf1Δ* strain expressing *RPL22B* under the control of the *GAL* promoter compared to the *upf1Δ* strain expressing *RPL22B* from its natural promoter and grown in galactose medium (Fig. 6, lanes 10 and 12). Thus, alternative splicing regulation of *RPL22B* upon heat shock is tightly linked to the identity of the *RPL22B* promoter, as switching the identity of the promoter is sufficient to favor the use of the normal 5′-splice site. The mechanism by which the identity of the

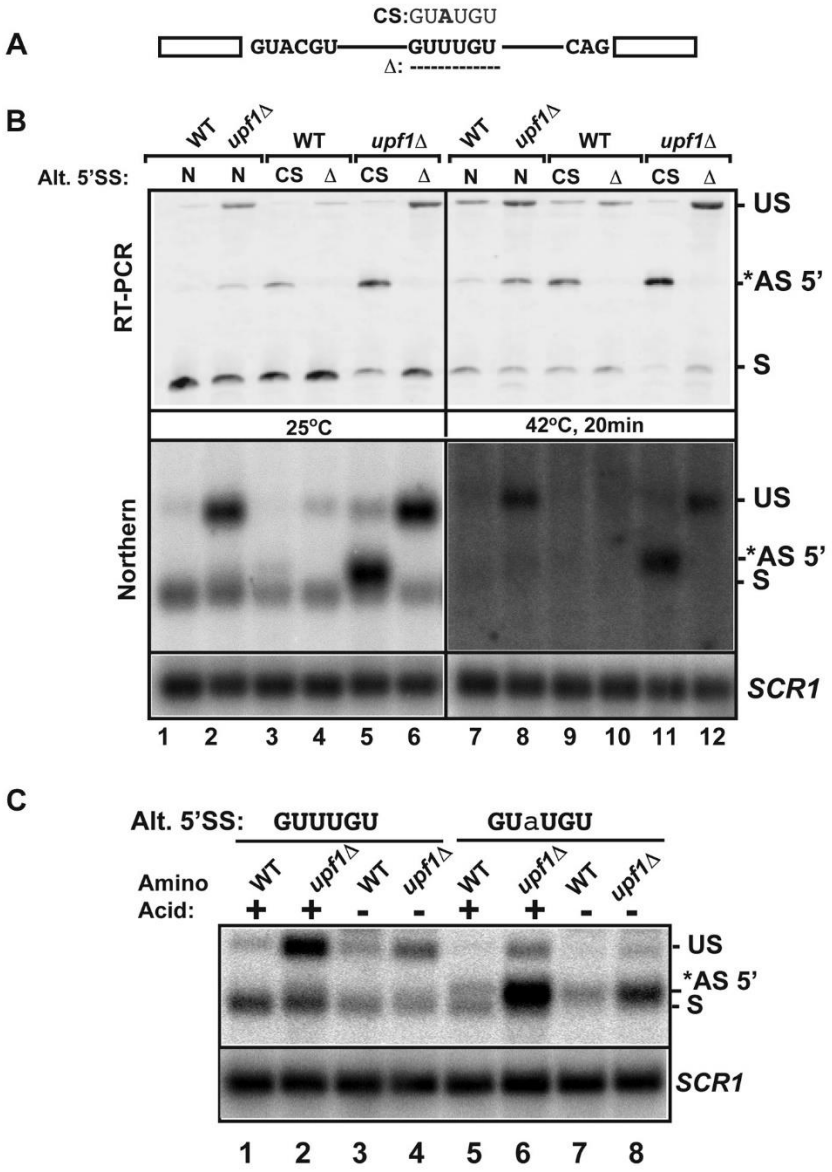


Figure 5. Effects of mutations of the *RPL22B* alternative 5′ splice site on *RPL22B* splicing and expression in normal and stress conditions. A. Organization of the *RPL22B* precursor, with the normal and alternative 5′-splice sites. Shown are the mutations to the consensus sequence (CS) GUAUGU or the deletion that entirely removes the GUUUGU sequence. B. Analysis of the effect of these mutations on *RPL22B* splicing and expression at normal temperatures (25°C) or after a 20 min heat shock at 42°C. N, natural 5′-splice site (GUUUGU); CS, consensus sequence (GUAUGU); Δ = deletion of the alternative 5′-splice site. Top panel: RT-PCR analysis. Bottom panel: northern blot analysis. US, *AS-5′, and S indicate the location of the products corresponding to the unspliced, alternatively spliced and normal spliced products, respectively. For the northern blot, *SCR1* was used as a loading control. C. Analysis of the effect of the *RPL22B* alternative splice site consensus mutation on *RPL22B* expression during amino acid starvation. Shown is a northern blot of RNA samples extracted from the indicated strains grown at 30°C in normal synthetic define complete (SDC) medium with amino acid (+) or in SDC medium lacking amino acid (−) for 10 minutes. Strains contained either the natural GUUUGU sequence at the alternative 5′-splice site of RPL22B, or the consensus GUaUGU sequence. The nucleotide mutated is highlighted in lower case. Labeling of the different species is similar to that of panel B. *SCR1* was used as a loading control.
doi:10.1371/journal.pgen.1004249.g005

promoter influences alternative splice site selection is unclear, but could be linked to the influence of the promoter on the speed of transcription. Nevertheless, we can conclude from these results that transcriptional down-regulation and the increased use of the alternative 5′-SS provide synergistic mechanisms to limit the expression of *RPL22B* during stress, consistent with the global down-regulation of ribosome biogenesis during stress conditions.

Discussion

A significant fraction of splicing events in *S. cerevisiae* generates non-functional RNA or protein products

In this study we show that the ensemble of transcripts generated by splicing from the *S. cerevisiae* genome is highly complex. Most of the splicing events that we have characterized in this study are

non-productive, either because they result in transcripts that are targeted by NMD, or because the protein products generated from these transcripts are predicted to be non-functional (e.g. *GPI15*). The large number of additional splice sites identified, and their relaxed conservation (Fig. 1D) imply that the rules governing splice selection are intrinsically more flexible than previously thought. This is further illustrated by the finding that a non-canonical AUG sequence in *GCR1* can be used as a 3′-SS in the absence of Prp18p (Fig. 3). In some cases, non-productive alternatively spliced transcripts accumulate only at low levels (e.g. *GCR1*, *GPI15*, Fig. 3). However, for other genes such as *TFC3*, the alternatively spliced non-productive transcripts represent a significant fraction (close to 30%) of all RNAs generated from this locus. Thus, non-productive splicing can significantly limit the expression of these genes. This was further demonstrated by mutagenesis of the non-productive splice site of *RPL22B*, as changing this site to a consensus sequence was sufficient to increase the splicing efficiency and the expression of this gene (Fig. 5B). Thus, the presence of alternative and sometimes sub-optimal splice sites that compete with the normal splice site contributes to an overall decrease in the amount of productively-spliced transcripts. Because the overlap in the alternative splicing events detected in all three NMD-deficient strains was limited (Fig. 1C), and because we detected by RT-PCR some alternative splicing events that escaped detection by RNA-Seq (e.g. *GCR1*), it is likely that we have not exhaustively identified the ensemble of splice sites that can be used by *S. cerevisiae*, and that additional splice sites will be identified by deeper sequencing or systematic RT-PCR analysis.

Contribution of splicing factors to alternative splice site selection and splice site fidelity

The analysis of double mutants in which splicing factor mutations were combined with NMD inactivation revealed some important and unexpected functions for these factors on alternative splice site selection. We found that the Nam8p and Mud1p components are important for the selection of some, but not all of the alternative 5′-splice sites described here. In the case of *RPL22B*, this requirement was likely due to the fact that the alternative 5′-SS possesses a suboptimal splicing sequence, and therefore exhibits a weaker affinity for U1 binding, and a stronger requirement for Mud1p and Nam8p that impact the efficiency of U1 snRNP assembly on the alternative splice site. Strikingly, we identified a new role for Prp17p in favoring the use of upstream, branchpoint-proximal 3′-SS. In all cases that we have analyzed, Prp17p inactivation resulted in an increase in the use of the downstream 3′-SS. The mechanistic basis for this novel function that we describe here for Prp17p in promoting branchpoint proximal 3′-SS is not fully understood. Because 3′-SS close to the branchpoint are often the first ones that are being used, this novel function for Prp17p could be linked to promoting the ability of the spliceosome to scan and recognize 3′-SS close to the branchpoint, or to unwind secondary structures that mask branchpoint-proximal 3′-SS. The absence of Prp17p would result in a higher rate of misrecognition of 3′-SS and in the use of more distal 3′-SS. In addition, we found that the absence of Prp18p resulted in the selection of a non-canonical AUG 3′-SS in *GCR1*, and that this atypical 3′-SS was utilized to a greater extent during heat-shock, revealing a unique function for Prp18p in suppressing usage of a

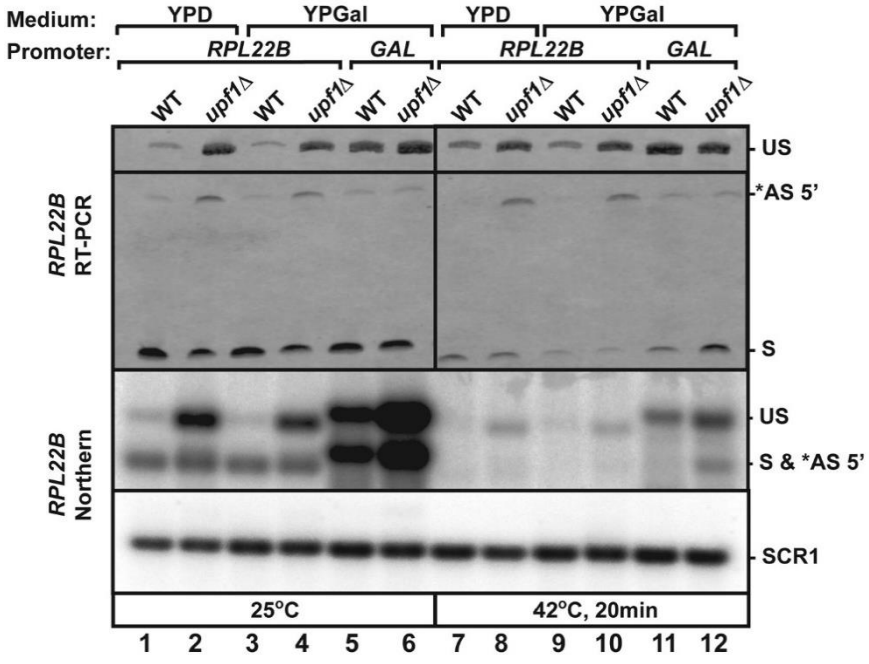


Figure 6. Replacement of *RPL22B* promoter by the GAL promoter results in a decrease in alternative 5′-splice site usage. Shown are the products generated when growing the indicated strains (wild-type or *upf1Δ* that contained the natural *RPL22B* promoter or the *GAL* promoter upstream *RPL22B*) in glucose (YPD) or galactose (YPGal)-containing media. Top panel, RT-PCR analysis. US, *AS 5′, and S indicate the location of the products corresponding to the unspliced, alternatively spliced and normal spliced products, respectively. Bottom Panel, Northern blot analysis. The labeling of the different species is similar to that of the top panel. SCR1 was used as a loading control.
doi:10.1371/journal.pgen.1004249.g006

non-canonical 3′-SS. This function for Prp18p is independent from Prp22p's function in proofreading 3′-SS [41], but might complement its role to ensure the overall proper fidelity of 3′-SS selection. While we have demonstrated this function for *GCR1* only, a full genomic analysis of 3′-SS usage in the absence of Prp18p might reveal further examples of non-canonical 3′-SS being used.

Spliceosome errors or bona-fide regulations?

The widespread occurrence of non-productive splice site usage described in this study begs the question of whether the use of these splice sites is the result of mistakes by the spliceosome, which occur at low frequency (as one might suggest based on their weaker consensus sequences) or whether they correspond to sites that have been selected throughout evolution for regulatory purposes. The sequence of some of these intronic, non-productive splice sites is conserved across closely related yeast species (Fig. S8 and S9), which, given the low conservation of intronic sequences in general, argues that this might reflect some degree of functional relevance. In addition, there is no obvious correlation between transcript levels and the occurrence of alternative splicing events (Fig. 1D), which argues against the suggestion that most of the alternative splicing events that we have mapped arise from low fidelity splicing events or errors that occur randomly, and which would be expected to be more frequently detected in highly abundant transcripts. Also, replacement of the *RPL22B* gene promoter results in higher transcript levels but reduces the usage of the alternative 5′-splice site of *RPL22B* (Fig. 6), providing another independent argument to suggest that the level of usage of alternative splice sites is not solely a reflection of overall transcript abundance. Finally, we show that the use of some of these alternative splice sites can be up-regulated during stress conditions (*RPL22B*, *GCR1*), and that this increased use participates in the down-regulation of *RPL22B* in stress conditions. Thus, this phylogenetically conserved, alternative, non-productive 5′-SS of *RPL22B* is functionally important because it contributes to the down-regulation of *RPL22B* during stress. This is shown by the fact that changing this sequence to a consensus sequence results in a significant increase in transcript levels upon NMD inactivation during stress (Fig. 5). The transcriptional down-regulation of ribosomal proteins during stress has been documented previously [45]. We show here that the promoter of the *RPL22B* gene is essential not only because it drives transcriptional repression during stress, but also because it controls the switch in 5′-SS selection that contributes to the overall repression of *RPL22B* during heat-shock. Thus, a combination of transcriptional and post-transcriptional regulations, through splicing inhibition [43,44], degradation of unspliced RNAs by NMD [12,47] and use of non-productive splice sites (this study) contributes to the repression of ribosomal protein production during stress. While several non-RPG transcripts analyzed in these stress conditions did not shown any changes, *GCR1* did exhibit a change in the use of alternative splice sites during stress (Fig. 4B). This result raises the possibility that other intron-containing genes may be regulated similarly by alternative splicing as a function of different environmental growth conditions. Overall our study has revealed that the pattern of splicing events in the model eukaryote *S. cerevisiae* is highly complex, but masked by NMD-mediated degradation. Given the recent report that another single cell eukaryote, *S.pombe* shows alternative splicing patterns conserved in higher eukaryotes [50], these observations suggest that alternative splicing provides an important contribution to genetic regulations and adaptations to environmental changes in unicellular eukaryotes. However, in the case of *S.cerevisiae*, the use of alternative splice sites have evolved towards fine tuning transcript levels, rather than generating proteome diversity as shown in higher eukaryotes.

Materials and Methods

Yeast culture and RNA analysis

Yeast strains were grown at 25°C in YPD medium, unless indicated otherwise in the figures. For heat shock treatment, strains were pre-grown in YPD at 25°C, spun down in 50 mL Falcon tubes, resuspended in pre-warmed YPD medium and heat shocked for 20 min before harvesting. For LiCl treatment, yeast strains were grown to mid-log phase in YPD rich media at 30°C, harvested by centrifugation in 50 mL Falcon tubes, washed once with pre-warmed 50 mL of YPD+300 mM LiCl before being resuspended in pre-warmed YPD with 300 mM LiCl for 10 minutes. For Rapamycin treatment, cells were grown to mid-log phase in rich media (at 30°C), and rapamycin from a stock solution of 1 mg/mL in 90% ethanol, 10% Tween-20 was added to a final concentration of 200 ng/mL and cells were incubated for 20 minutes. The same volume of 90% ethanol, 10% Tween-20 solution used for the rapamycin treatment was added to the negative control. Sample preparation and RNA sequencing was performed by Illumina. RT-PCR analysis and northern blot was performed as described [21].

Mapping reads

High throughput sequencing data have been deposited in the GEO database (accession GSE55213). All sequence files were aligned against the 2008 SGD assembly of the *Saccharomyces cerevisiae* genome. The novoalign software package (www.novocraft.com) and the BLAT alignment tool [22] were used to align 75 base pair reads in two steps. In the first step, sequences were aligned with novoalign allowing for up to four mismatches and no gaps. In the second step, sequences that failed to align in the first step were aligned with BLAT allowing three mismatches and gaps up to 20000 nucleotides in length. A sequence was kept for further analysis if it mapped with equal score to at most two genomic locations and did not contain a gap smaller than ten nucleotides.

Intronic sequences counts

Intronic sequence expression representative of unspliced RNAs was quantified for each ICG by summing reads that aligned to introns and exon-intron boundaries. Values between samples were normalized by total mapped reads to account for lane effects. p-values were computed by modeling each ICG wild-type count as a poisson random variable and calculating the probability of observing each mutant count if it were drawn from the same distribution.

Quantification of alternative splicing events

Alternative splicing events were defined as splicing events that are within ICGs and are supported by sequencing but that are not annotated in the *Saccharomyces* Genome Database (SGD). Counts of total alternative splicing events and PTC-generating alternative splicing events were quantified by summing all unique alternative splicing events in each sample. To determine if an alternative splicing event is PTC-generating we constructed the splice product's sequence using the novel splicing event in the otherwise canonical transcript sequence. Counts between samples were normalized by sequencing depth. p-values were calculated by modeling the wild-type count as a poisson random variable and calculating the probability of observing each mutant's count for both total alternative splicing events and PTC-generating

alternative splicing events. Venn diagrams of agreement between samples were generated using BioVenn [51].

Splice site consensus sequence

Consensus sequences for 5′ and 3′ ends of both canonical splice sites and alternative splice sites were represented as sequence logos. Sequence logos were constructed using the MATLAB (MathWorks) seqlogo function.

Supporting Information

Figure S1 Quantification of the SRC1-L and SRC1-S isoforms in wild-type, *upf1Δ* and splicing mutants. Shown is the percentage of the *SRC1-L* and *SRC1-S* transcripts in various strains. Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S2 Quantification of the usage of the normal and alternative 5′-splice sites of *RPL22B* in wild-type, *upf1Δ* and splicing mutants. Shown is the percentage of transcripts spliced at the normal 5′-splice site (red) and at the alternative 5′-splice site (blue). Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S3 Quantification of the usage of the normal and alternative 3′-splice sites of *TFC3* in wild-type, *upf1Δ* and splicing mutants. Shown is the percentage of transcripts spliced at the normal 3′-splice site (blue) and at the alternative 5′-splice site (red). Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S4 Quantification of the usage of the two alternative 3′-splice sites of *TAN1* in wild-type, *upf1Δ* and splicing mutants. Shown is the percentage of transcripts spliced at the alternative 3′-splice site #1 (blue) or #2 (red) compared to all the spliced transcripts. Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S5 Validation of the use of the AUG alternative 3′ splice site of *GCR1* by RT-PCR. Sequencing of the cloned *D and *E cDNAs determined the location of the splice junction, while sequencing of unspliced cDNAs was used to confirm that this unusual alternative 3′-SS was indeed AUG, and not a SNP or other mutation of the *GCR1* gene that would have converted it into an AAG. RT-PCR confirmation of the use of this AUG 3′-SS was performed using reverse primers spanning the splice junction to specifically amplify distinct splicing events; either associated with *D, *E, or unspliced. The use of the AUG 3′ SS was also confirmed using an intronic reverse primer just downstream of the AUG sequence and detected *D, *E, and unspliced products, as predicted (Fig. S1). A. RT-PCR strategy. All PCR include the same forward primer For, and various reverse primers that hybridize to the indicated regions of *GCR1*. B. RT-PCR data. Shown are the PCR products obtained from the different reverse primers shown in A.
(TIF)

Figure S6 Quantification of the abundance of the major alternatively spliced forms of *GCR1* in wild-type, *upf1Δ* and splicing mutants. Shown is the percentage of the *D (blue), *A (red) or *C (green) spliced forms. Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S7 RT-PCR analysis of GCR1 splicing in the *prp18* and *prp22-1* mutant strains. The identity of the different spliced products is labeled according to Figure 2.
(TIF)

Figure S8 Conservation of the intronic alternative 5′-SS in *RPL22B*. **A**. Screen capture of the web browser showing the RNA-Seq reads mapped for *RPL22B* that use the alternative 5′-SS in black, and the sequence conservation in closely related yeast species as blue peaks. The peak showing conservation of the intronic alternative 5′splice site is shown on the right, since the gene is encoded on the Crick strand. **B**. Zoomed in view of the conservation of the sequence of the alternative 5′-SS (ACAAC sequence because of the Crick Strand).
(PDF)

Figure S9 Conservation of the intronic alternative 3′-SS in *TAN1*. **A**. Screen capture of the web browser showing the RNA-Seq reads mapped for *TAN1* that use the intronic alternative 3′-SS in black, and the sequence conservation in closely related yeast species as blue peaks. **B**. Zoomed in view of the conservation of the sequence of the alternative intronic AAG 3′-SS.
(PDF)

Figure S10 Quantification of the usage of the alternative 5′-splice sites of *RPL22B* in various normal media (YPD, SDC) or in stress conditions (Heat shock, amino acid starvation). Values shown are the average and standard deviations obtained from RT-PCR experiments of three independent cultures for each strain.
(TIF)

Figure S11 RT-PCR analysis of the spliced products of *TAN1* and *TFC3* under stress conditions. Shown are the products for the unspliced (US), normal spliced product (S), and the alternatively spliced species (AS) described in Figure 2.
(TIF)

Figure S12 Quantitation of the use of the alternative 5′-splice site of *RPL22B* under normal growth conditions (25°C) and after a 20 min heat shock at 42°C. Plotted are the amount of transcript spliced at the alternative splice site divided by the values obtained for all spliced species for the indicated strains. Shown are the average of 4 to 5 independent experiments with the standard deviations.
(TIF)

Figure S13 Quantitation of the use of the normal and alternative 5′-splice site of *RPL22B* under normal growth conditions in minimal medium (SDC) and after amino acid starvation (-AA) for the strains expressing the natural (N) GUUUGU sequence at the alternative 5′ splice site of *RPL22B*, or the consensus (CS) GUAUGU sequence in the context of wild-type *UPF1* (WT) or when *UPF1* has been deleted (**Δ**). Plotted are the amount of transcript spliced at the normal and alternative splice sites divided by the values obtained for all spliced species. Shown are the average of 3 independent experiments with the standard deviations.
(TIF)

Table S1 Statistics of RNA-Seq analysis sequence alignments.
(XLSX)

Table S2 Number of alternative splicing events detected in wild-type and NMD-deficient strains.
(XLSX)

Table S3 Mapping of RNA-Seq reads in wild-type and NMD-deficient strains in various genomic elements and in intron-containing genes.
(XLSX)

Table S4 List of intron-containing genes for which no alternative splicing events were detected. Shown is the list of intron-containing genes for which no alternative splicing junctions were detected in any of the strains sequenced. The adjusted number of reads obtained from each strain (RPKM) and the sequence of the 5′- and 3′-splice sites is shown for each of these genes.
(XLSX)

Table S5 List of intron-containing genes for which alternative splicing events were detected. Shown is the list of intron-containing genes for which alternative splicing junctions were detected in at least one of the strains sequenced. Shown is the number of splice junction sequences counts for the normal (blue) and alternative (green) splicing events. For the alternative splicing events, the 5′ and/or 3′ splice sites which differ from the normal splice sites are marked by an asterisk. The position of these splice sites on each chromosome is also indicated. The adjusted number of reads obtained from each strain (RPKM) and the sequence of the normal and alternative 5′- and 3′-splice sites is shown for each of these genes.
(XLSX)

Table S6 List of alternative proteins potentially generated by alternative splicing in wild-type or NMD mutants. For each open reading frame, a portion of the normal protein sequence is shown on the first line, and the sequence that differs upon the alternative splicing event is shown below. Amino acids maintained between the two forms are indicated in red. Amino acids that differ between the two forms are highlighted in bold and black. In the case of splicing events inducing a deletion, a delta sign has been added with a number corresponding to the number of amino acid deleted. The numbers of the first and last amino acids shown is indicated before and after each protein sequence, respectively. The numbers in brackets that follow each protein sequence correspond to the number of reads for the splice junctions in the wild-type strain, *upf1Δ*, *upf2Δ* and *upf3Δ* mutants.
(DOCX)

Acknowledgments

We thank J. Staley for insights, discussions and comments on the manuscript and K. Roy for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: TK SD JG MP GFC. Performed the experiments: TK JG. Analyzed the data: TK SD JG MP GFC. Wrote the paper: TK GFC.

References

1. Isken O, Maquat LE (2007) Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev 21: 1833–1856.
2. Kervestin S, Jacobson A (2012) NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol 13: 700–712.
3. Green RE, Lewis BP, Hillman RT, Blanchette M, Lareau LF, et al. (2003) Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics 19 Suppl 1: i118–121.
4. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446: 926–929.
5. Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, et al. (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21: 708–718.
6. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet 36: 1073–1078.
7. Chan WK, Huang L, Gudikote JP, Chang YF, Imam JS, et al. (2007) An alternative branch of the nonsense-mediated decay pathway. EMBO J 26: 1820–1830.
8. Weischenfeldt J, Damgaard I, Bryder D, Theilgaard-Monch K, Thoren LA, et al. (2008) NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev 22: 1381–1396.
9. He F, Peltz SW, Donahue JL, Rosbash M, Jacobson A (1993) Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proc Natl Acad Sci U S A 90: 7034–7038.
10. Mitrovich QM, Anderson P (2000) Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes Dev 14: 2173–2184.
11. Jaillon O, Bouhouche K, Gout JF, Aury JM, Noel B, et al. (2008) Translational control of intron splicing in eukaryotes. Nature 451: 359–362.
12. Sayani S, Janis M, Lee CY, Toesca I, Chanfreau GF (2008) Widespread impact of nonsense-mediated mRNA decay on the yeast intronome. Mol Cell 31: 360–370.
13. Sayani S, Chanfreau GF (2012) Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in Saccharomyces cerevisiae. RNA 18: 1563–1572.
14. Vijayraghavan U, Company M, Abelson J(1989) Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae. Genes Dev 3: 1206–1216.
15. Wahl MC, Will CL, Luhrmann R (2009) The spliceosome: design principles of a dynamic RNP machine. Cell 136: 701–718.
16. Grund SE, Fischer T, Cabal GG, Antunez O, Perez-Ortin JE, et al. (2008) The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. J Cell Biol 182: 897–910.
17. Mishra SK, Ammon T, Popowicz GM, Krajewski M, Nagel RJ, et al. (2011) Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature 474: 173–178.
18. Meyer M, Plass M, Perez-Valle J, Eyras E, Vilardell J (2011) Deciphering 3′ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. Mol Cell 43: 1033–1039.
19. Plass M, Codony-Servat C, Ferreira PG, Vilardell J, Eyras E (2012) RNA secondary structure mediates alternative 3′ss selection in Saccharomyces cerevisiae. RNA 18: 1103–1115.
20. Marshall AN, Montealegre MC, Jimenez-Lopez C, Lorenz MC, van Hoof A (2013) Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. PLoS Genet 9: e1003376.
21. Kawashima T, Pellegrini M, Chanfreau GF (2009) Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. RNA 15: 2236–2247.
22. Kent WJ (2002) BLAT–the BLAST-like alignment tool. Genome Res 12: 656–664.
23. Liao XC, Tang J, Rosbash M (1993) An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. Genes Dev 7: 419–428.
24. Neubauer G, Gottschalk A, Fabrizio P, Seraphin B, Luhrmann R, et al. (1997) Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A 94: 385–390.
25. Puig O, Gottschalk A, Fabrizio P, Seraphin B (1999) Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection. Genes Dev 13: 569–580.
26. Gottschalk A, Tang J, Puig O, Salgado J, Neubauer G, et al. (1998) A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. RNA 4: 374–393.
27. Umen JG, Guthrie C (1995) Prp16p, Slu7p, and Prp8p interact with the 3′ splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. RNA 1: 584–597.
28. Aronova A, Bacíková D, Crotti LB, Horowitz DS, Schwer B (2007) Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing. RNA 13: 1437–1444.
29. Villa T, Guthrie C (2005) The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes Dev 19: 1894–1904.
30. Saha D, Banerjee S, Bashir S, Vijayraghavan U (2012) Context dependent splicing functions of Bud31/Ycr063w define its role in budding and cell cycle progression. Biochem Biophys Res Commun 424: 579–585.
31. Rodriguez-Navarro S, Igual JC, Perez-Ortin JE (2002) SRC1: an intron-containing yeast gene involved in sister chromatid segregation. Yeast 19: 43–54.
32. Chanfreau GF (2010) A dual role for RNA splicing signals. EMBO Rep 11: 720–721.

33. Crotti LB, Horowitz DS (2009) Exon sequences at the splice junctions affect splicing fidelity and alternative splicing. Proc Natl Acad Sci U S A 106: 18954–18959.
34. Zhou Y, Chen C, Johansson MJ (2013) The pre-mRNA retention and splicing complex controls tRNA maturation by promoting TAN1 expression. Nucleic Acids Res 41: 5669–5678.
35. Yan BC, Westfall BA, Orlean P (2001) Ynl038wp (Gpi15p) is the Saccharomyces cerevisiae homologue of human Pig-Hp and participates in the first step in glycosylphosphatidylinositol assembly. Yeast 18: 1383–1389.
36. Muhlrad D, Parker R (1999) Aberrant mRNAs with extended 3′ UTRs are substrates for rapid degradation by mRNA surveillance. RNA 5: 1299–1307.
37. Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, et al. (2004) A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature 432: 112–118.
38. Uemura H, Jigami Y (1995) Mutations in GCR1, a transcriptional activator of Saccharomyces cerevisiae glycolytic genes, function as suppressors of gcr2 mutations. Genetics 139: 511–521.
39. Holland MJ, Yokoi T, Holland JP, Myambo K, Innis MA (1987) The GCR1 gene encodes a positive transcriptional regulator of the enolase and glyceraldehyde-3-phosphate dehydrogenase gene families in Saccharomyces cerevisiae. Mol Cell Biol 7: 813–820.
40. Clifton D, Fraenkel DG (1981) The gcr (glycolysis regulation) mutation of Saccharomyces cerevisiae. J Biol Chem 256: 13074–13078.
41. Mayas RM, Maita H, Staley JP (2006) Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol 13: 482–490.
42. James SA, Turner W, Schwer B (2002) How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. RNA 8: 1068–1077.
43. Bergkessel M, Whitworth GB, Guthrie C (2011) Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA 17: 1461–1478.
44. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C (2007) Rapid, transcript-specific changes in splicing in response to environmental stress. Mol Cell 27: 928–937.
45. Li B, Nierras CR, Warner JR (1999) Transcriptional elements involved in the repression of ribosomal protein synthesis. Mol Cell Biol 19: 5393–5404.
46. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, et al. (2000) Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11: 4241–4257.
47. Garre E, Romero-Santacreu L, Barneo-Munoz M, Miguel A, Perez-Ortin JE, et al. (2013) Nonsense-mediated mRNA decay controls the changes in yeast ribosomal protein pre-mRNAs levels upon osmotic stress. PLoS One 8: e61240.
48. Yost HJ, Lindquist S (1991) Heat shock proteins affect RNA processing during the heat shock response of Saccharomyces cerevisiae. Mol Cell Biol 11: 1062–1068.
49. Vogel JL, Parsell DA, Lindquist S (1995) Heat-shock proteins Hsp104 and Hsp70 reactivate mRNA splicing after heat inactivation. Curr Biol 5: 306–317.
50. Awan AR, Manfredo A, Pleiss JA (2013) Lariat sequencing in a unicellular yeast identifies regulated alternative splicing of exons that are evolutionarily conserved with humans. Proc Natl Acad Sci U S A 110: 12762–12767
51. Hulsen T, de Vlieg J, Alkema W (2008) BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9: 488.

Supporting Information

Figure S1

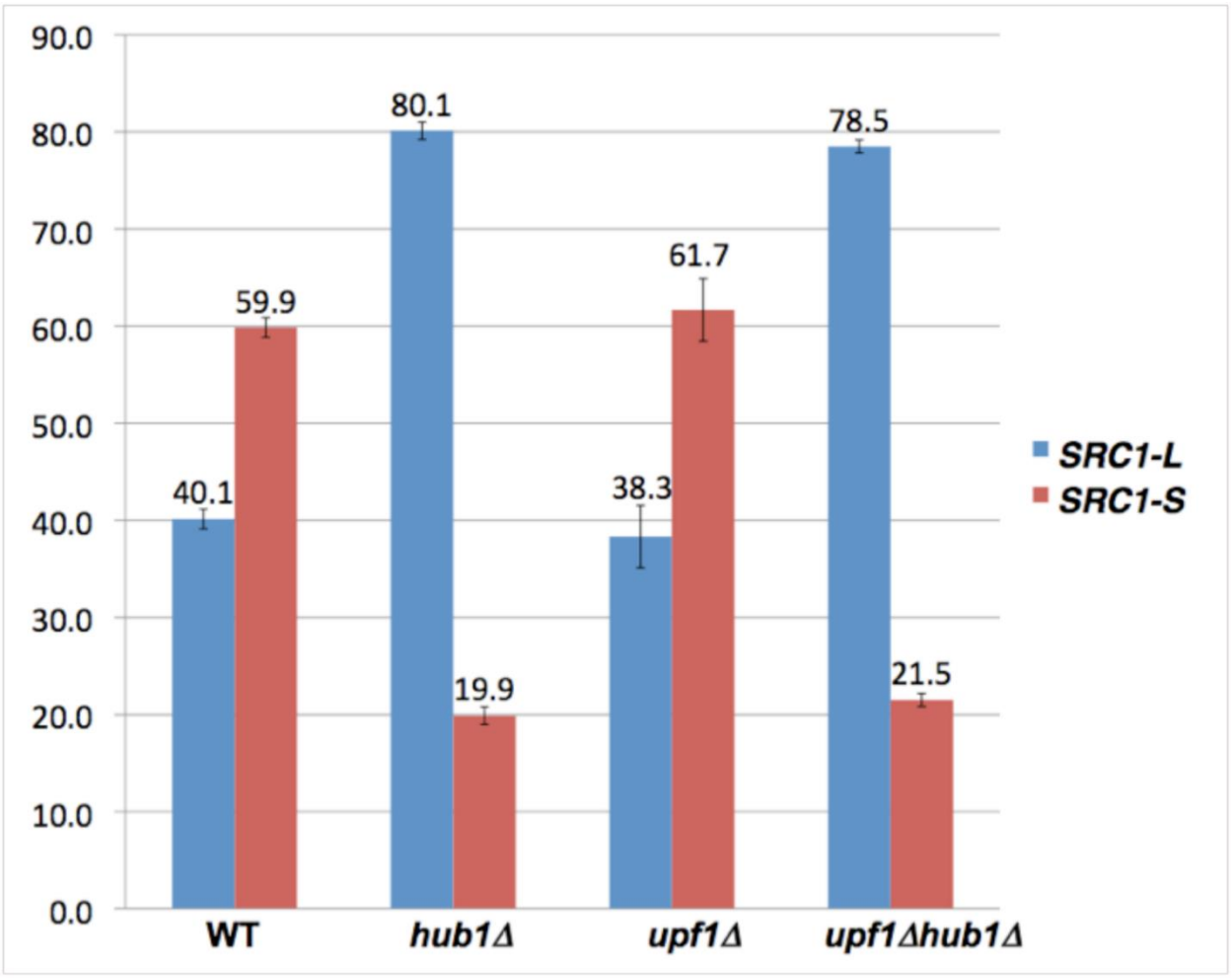


Figure S2

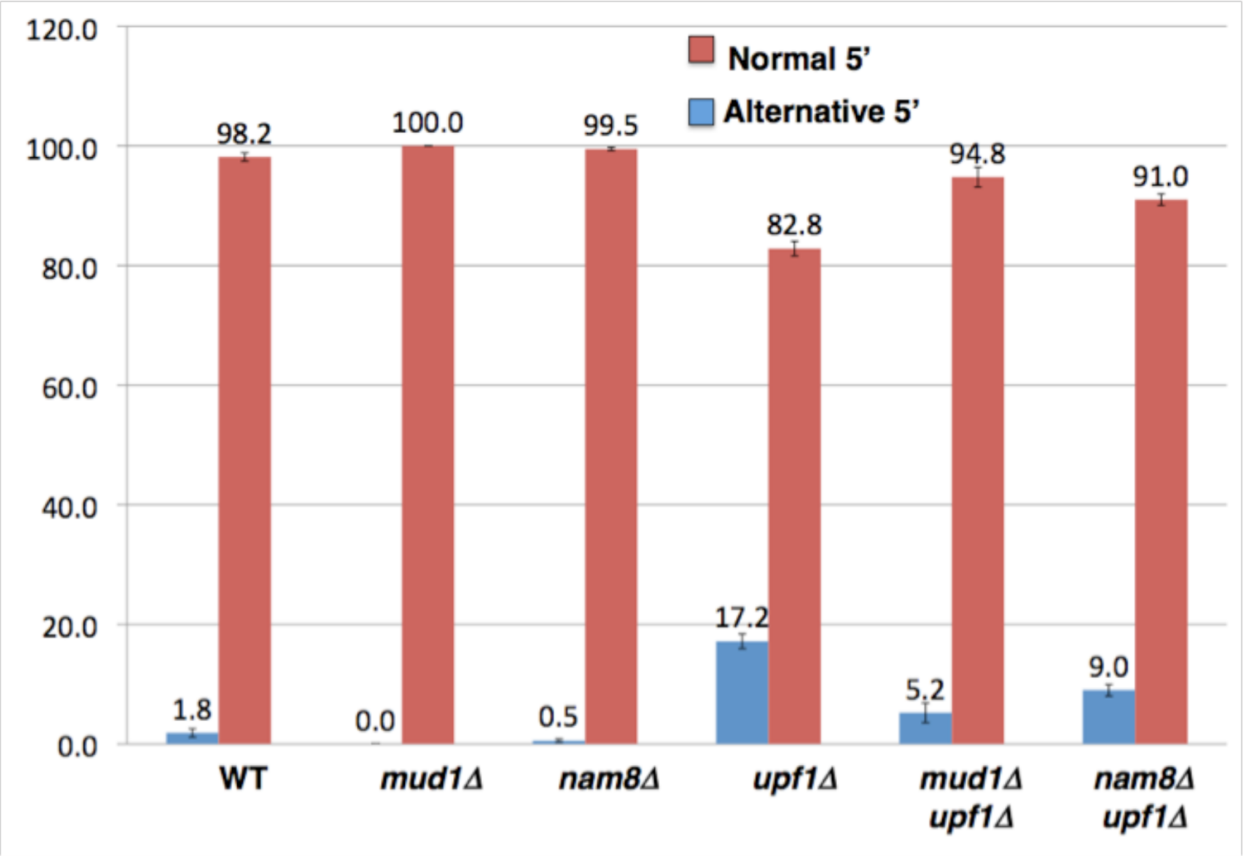

120.0
100.0
80.0
60.0
40.0
20.0
0.0
Normal 5'
Alternative 5'
98.2
100.0
99.5
82.8
94.8
91.0
1.8
0.0
0.5
17.2
5.2
9.0
WT
mud1Δ
nam8Δ
upf1Δ
mud1Δ upf1Δ
nam8Δ upf1Δ

Figure S3

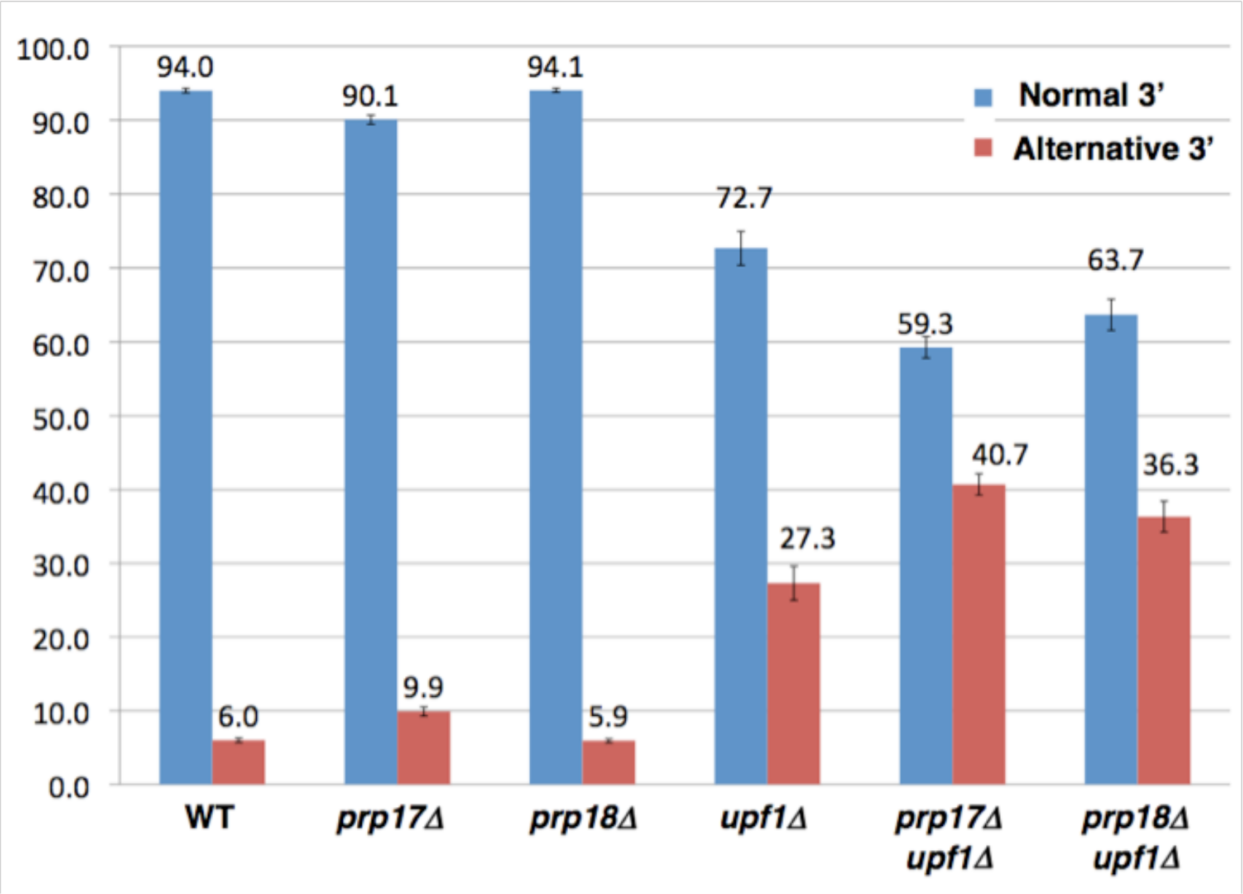

100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
94.0
6.0
90.1
9.9
94.1
5.9
72.7
27.3
59.3
40.7
63.7
36.3
Normal 3'
Alternative 3'
WT
prp17Δ
prp18Δ
upf1Δ
prp17Δ upf1Δ
prp18Δ upf1Δ

Figure S4

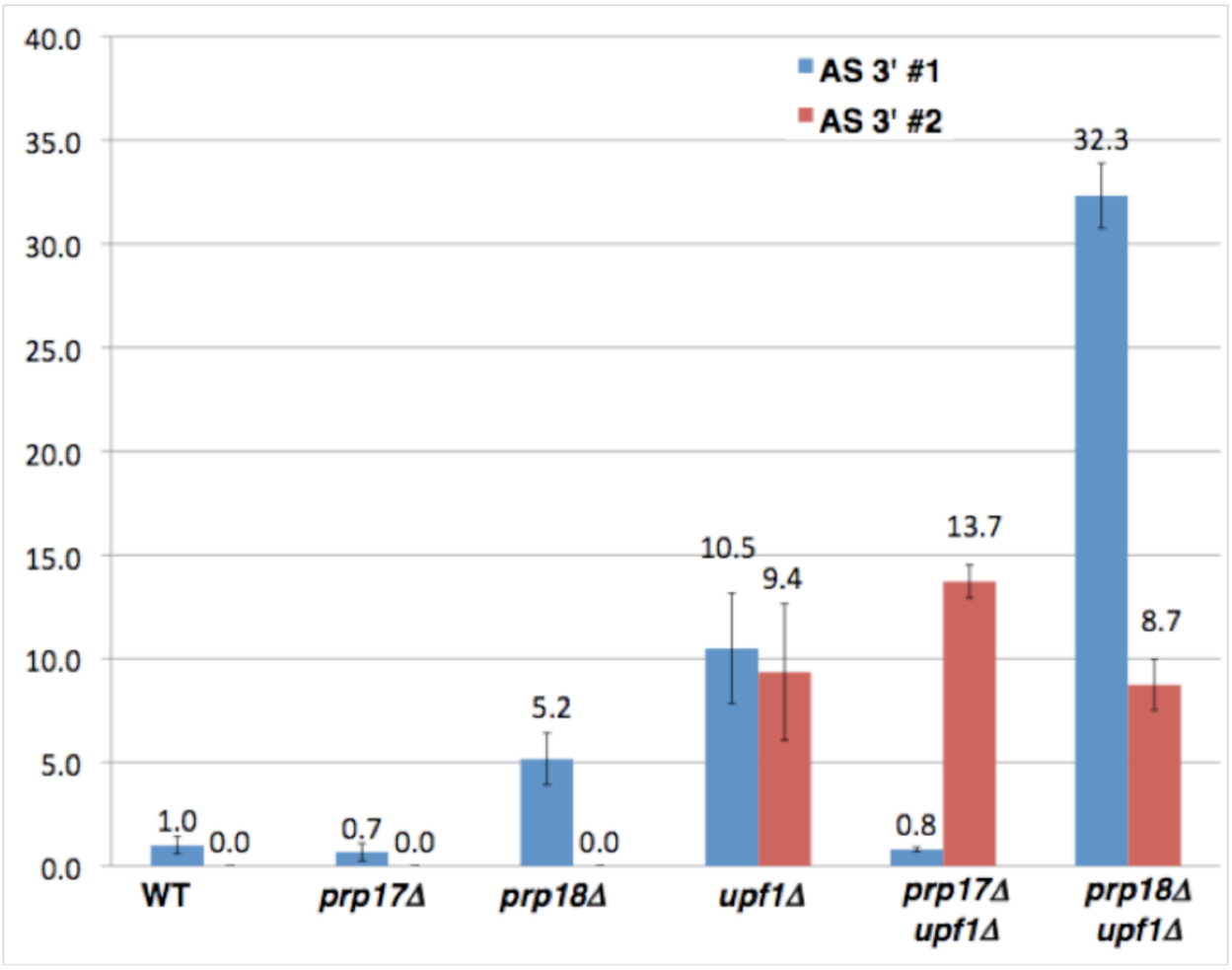

40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
AS 3' #1
AS 3' #2
1.0
0.0
0.7
0.0
5.2
0.0
10.5
9.4
0.8
13.7
32.3
8.7
WT
prp17Δ
prp18Δ
upf1Δ
prp17Δ upf1Δ
prp18Δ upf1Δ

Figure S5

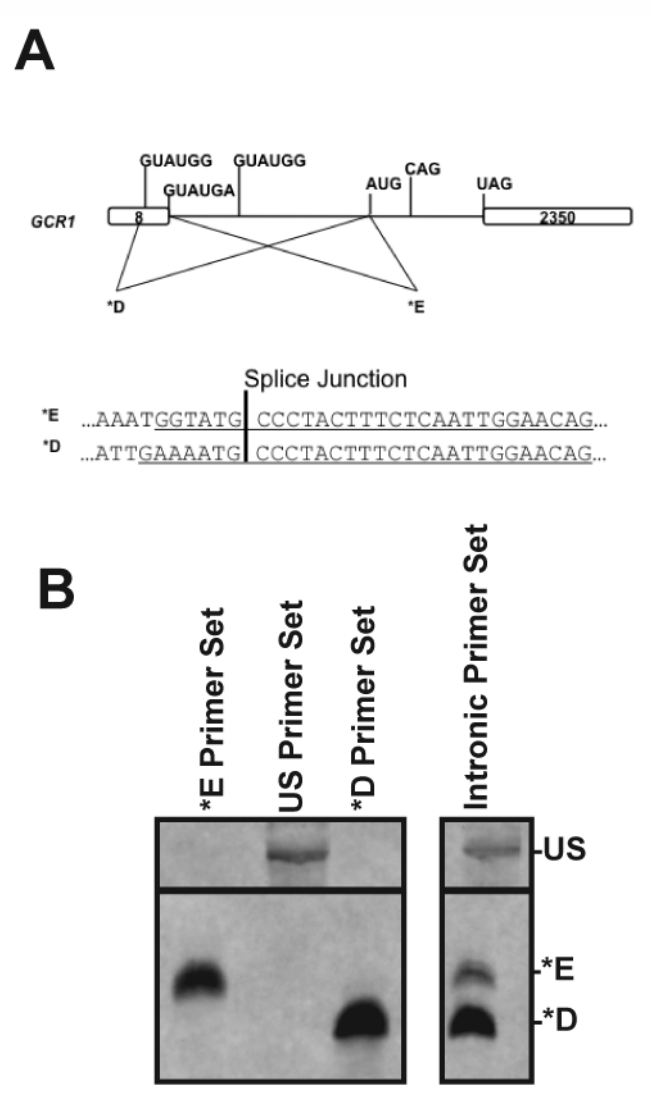

A
GUAUGG
GUAUGG
GUAUGA
AUG
CAG
UAG
GCR1
8
2350
*D
*E
Splice Junction
*E ...AAATGGTATG CCCTACTTTCTCAATTGGAACAG...
*D ...ATTGAAAATG CCCTACTTTCTCAATTGGAACAG...
B
*E Primer Set
US Primer Set
*D Primer Set
Intronic Primer Set
US
*E
*D

Figure S6

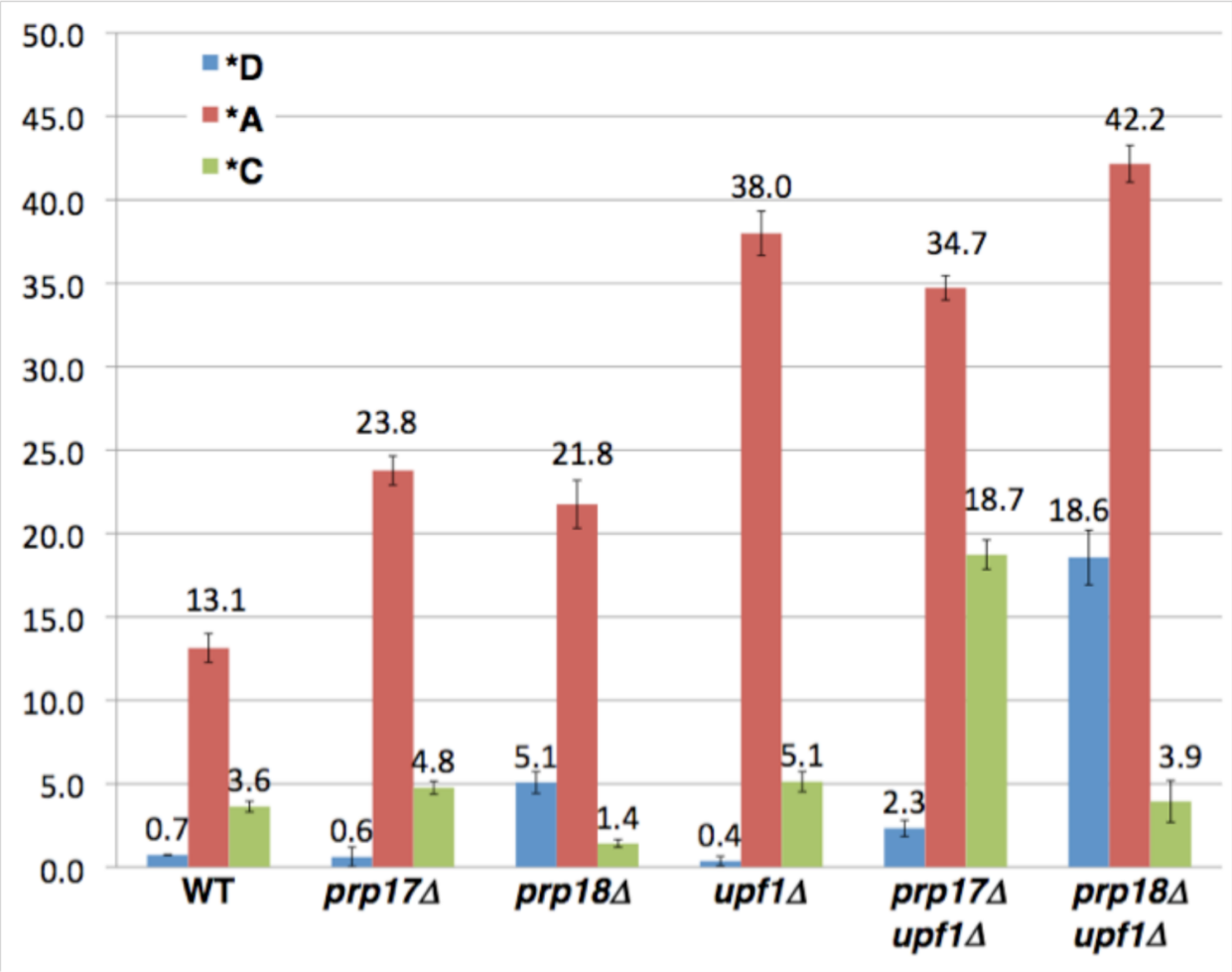

50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
*D
*A
*C
0.7
13.1
3.6
0.6
23.8
4.8
5.1
21.8
1.4
0.4
38.0
5.1
2.3
34.7
18.7
18.6
42.2
3.9
WT
prp17Δ
prp18Δ
upf1Δ
prp17Δ upf1Δ
prp18Δ upf1Δ

Figure S7

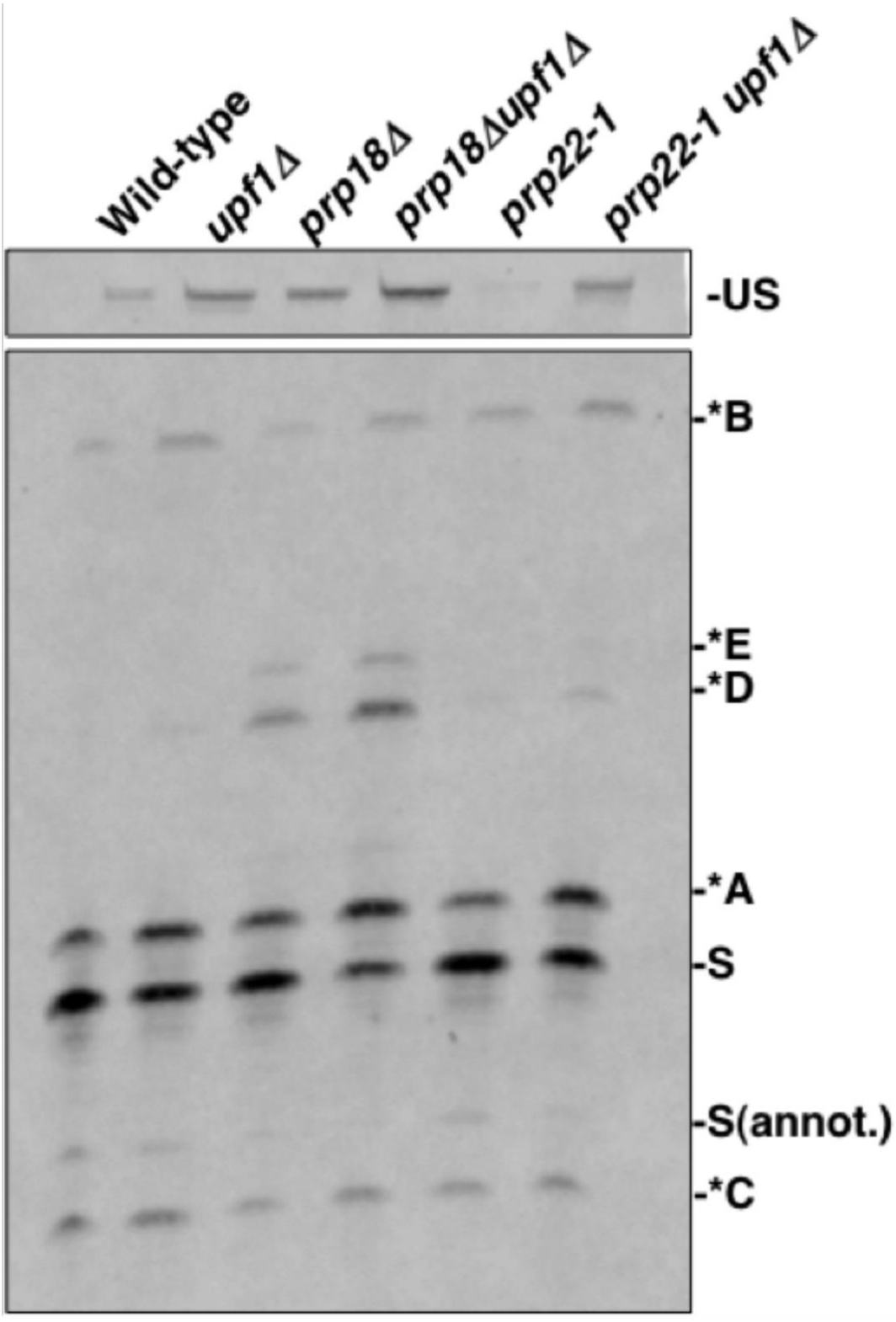
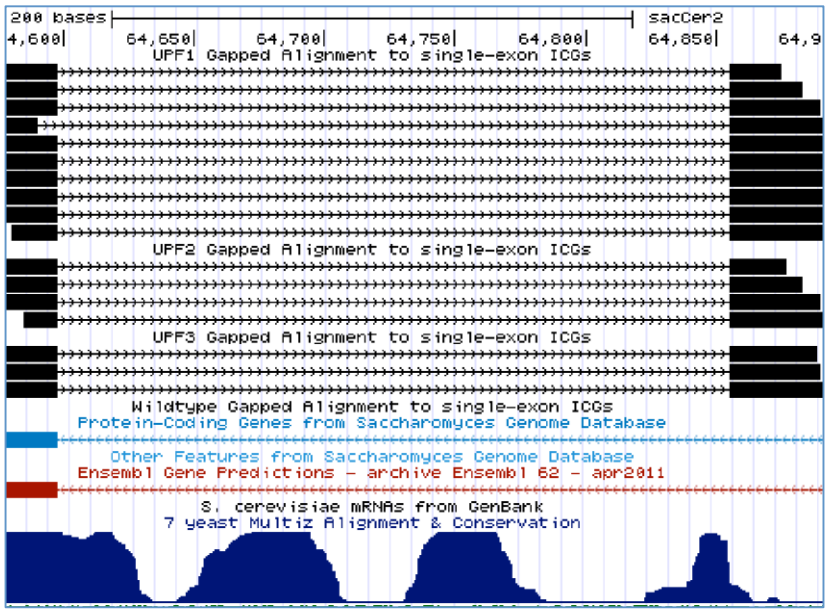

Wild-type
upf1Δ
prp18Δ
prp18Δupf1Δ
prp22-1
prp22-1 upf1Δ
-US
-*B
-*E
-*D
-*A
-S
-S(annot.)
-*C

Figure S8

A


200 bases
sacCer2
4,600|
64,650|
64,700|
64,750|
64,800|
64,850|
64,9
UPF1 Gapped Alignment to single-exon ICGs
UPF2 Gapped Alignment to single-exon ICGs
UPF3 Gapped Alignment to single-exon ICGs
Wildtype Gapped Alignment to single-exon ICGs
Protein-Coding Genes from Saccharomyces Genome Database
Other Features from Saccharomyces Genome Database
Ensembl Gene Predictions - archive Ensembl 62 - apr2011
S. cerevisiae mRNAs from GenBank
7 yeast Multiz Alignment & Conservation

B

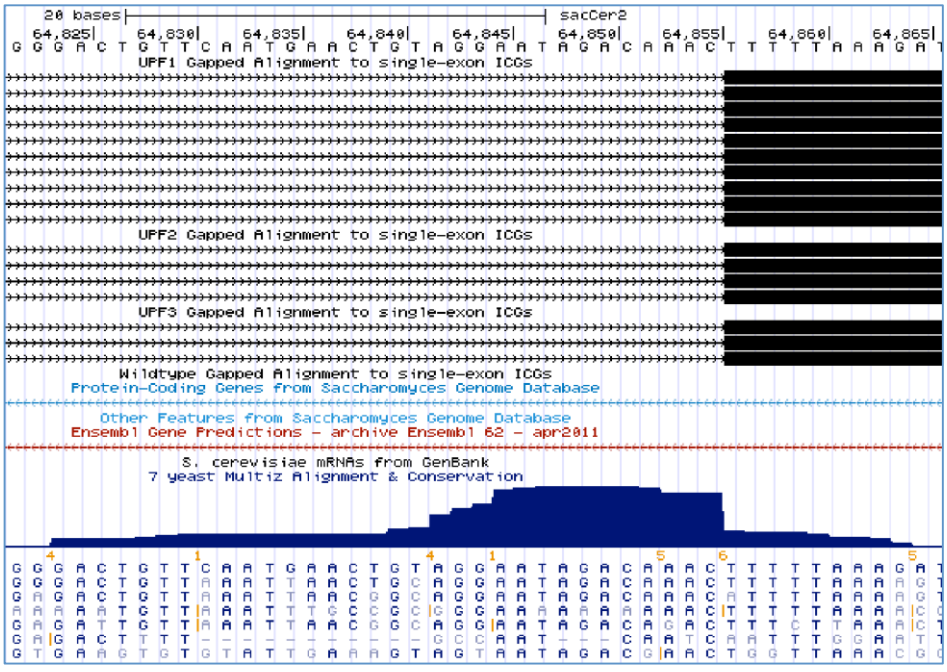
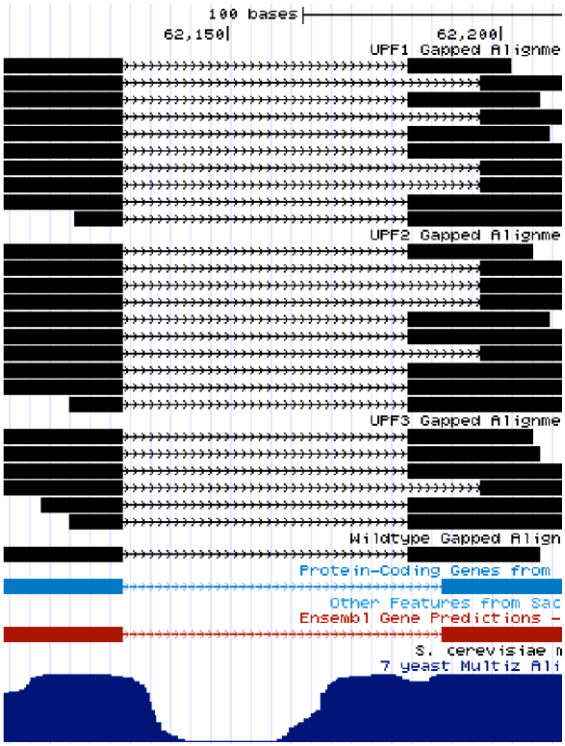

20 bases
sacCer2
64,825|
64,830|
64,835|
64,840|
64,845|
64,850|
64,855|
64,860|
64,865|
UPF1 Gapped Alignment to single-exon ICGs
UPF2 Gapped Alignment to single-exon ICGs
UPF3 Gapped Alignment to single-exon ICGs
Wildtype Gapped Alignment to single-exon ICGs
Protein-Coding Genes from Saccharomyces Genome Database
Other Features from Saccharomyces Genome Database
Ensembl Gene Predictions - archive Ensembl 62 - apr2011
S. cerevisiae mRNAs from GenBank
7 yeast Multiz Alignment & Conservation

Figure S9

A


100 bases
62,150
62,200
UPF1 Gapped Alignme
UPF2 Gapped Alignme
UPF3 Gapped Alignme
Wildtype Gapped Align
Protein-Coding Genes from
Other Features from Sac
Ensembl Gene Predictions –
S. cerevisiae n
7 yeast Multiz Ali

B

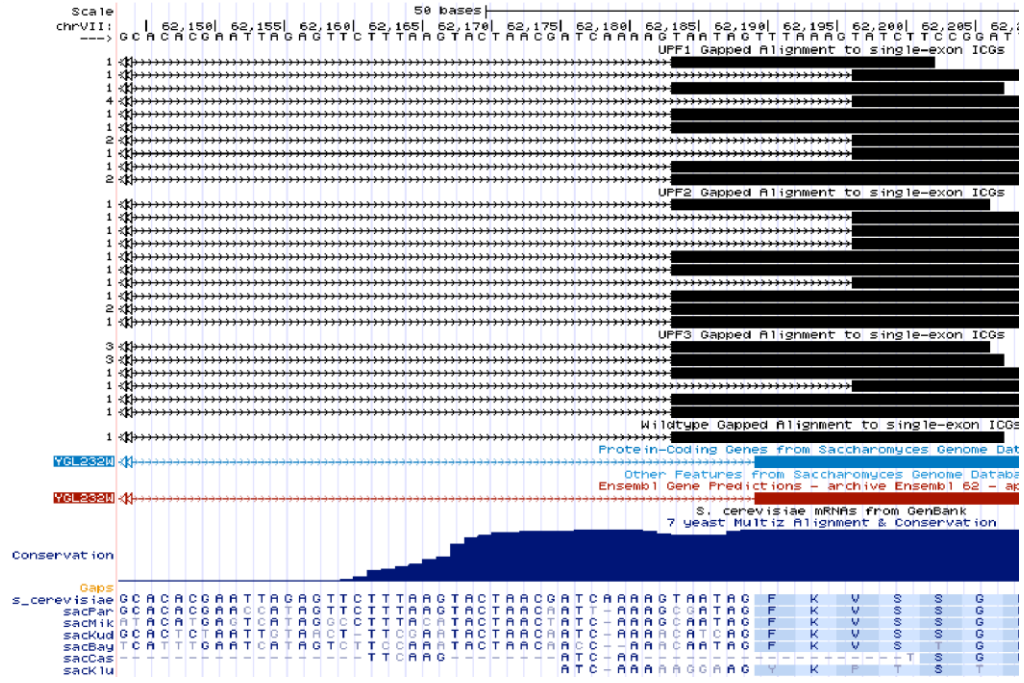

Scale
50 bases
chrVII:
UPF1 Gapped Alignment to single-exon ICGs
UPF2 Gapped Alignment to single-exon ICGs
UPF3 Gapped Alignment to single-exon ICGs
Wildtype Gapped Alignment to single-exon ICGs
Protein-Coding Genes from Saccharomyces Genome Dat
YGL232W
Other Features from Saccharomyces Genome Databa
Ensembl Gene Predictions – archive Ensembl 62 – ap
YGL232W
S. cerevisiae mRNAs from GenBank
7 yeast Multiz Alignment & Conservation
Conservation
Gaps
s_cerevisiae
sacPar
sacMik
sacKud
sacBay
sacCas
sacKlu

Figure S10

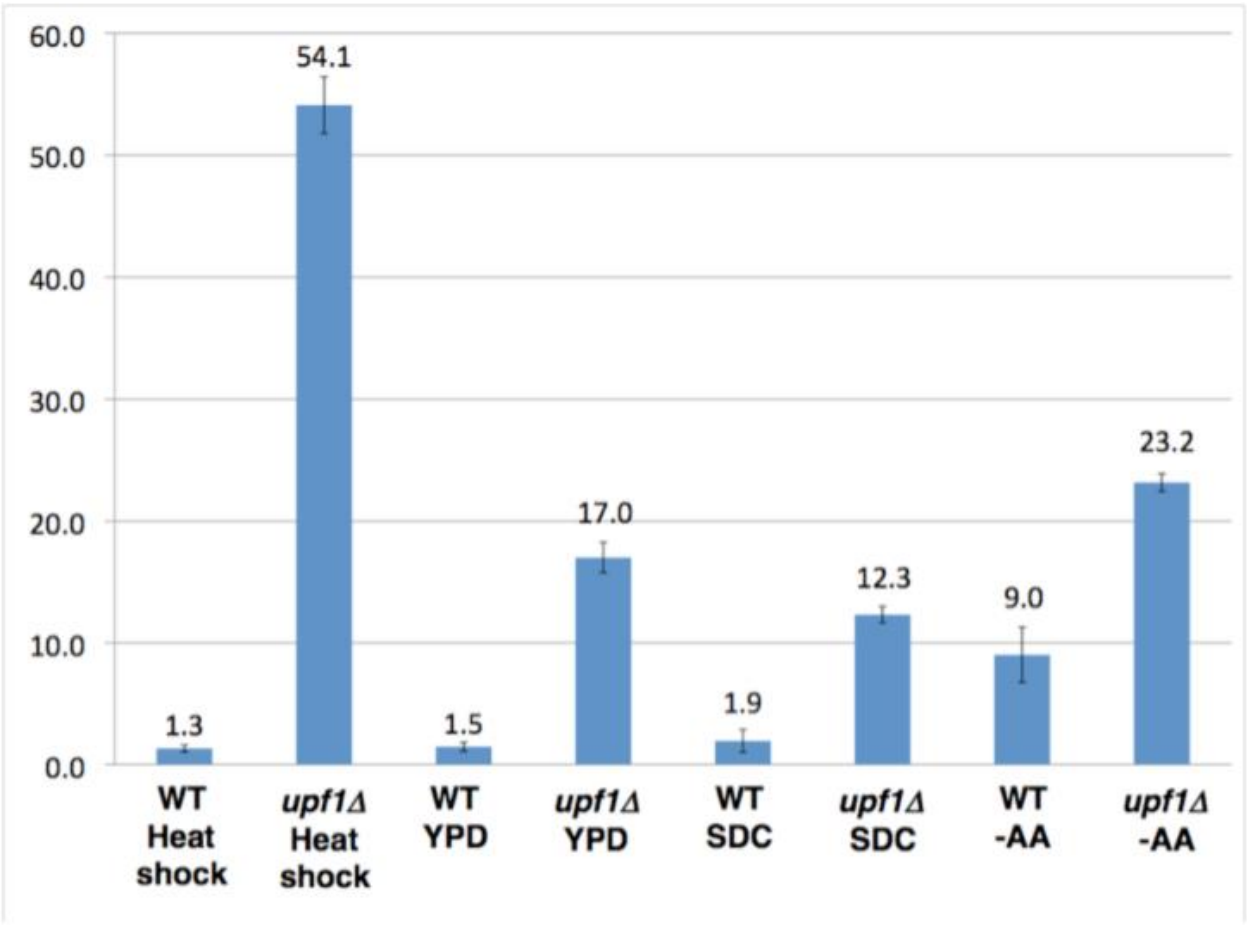

60.0
50.0
40.0
30.0
20.0
10.0
0.0
54.1
17.0
23.2
12.3
9.0
1.3
1.5
1.9
WT Heat shock
upf1Δ Heat shock
WT YPD
upf1Δ YPD
WT SDC
upf1Δ SDC
WT -AA
upf1Δ -AA

Figure S11

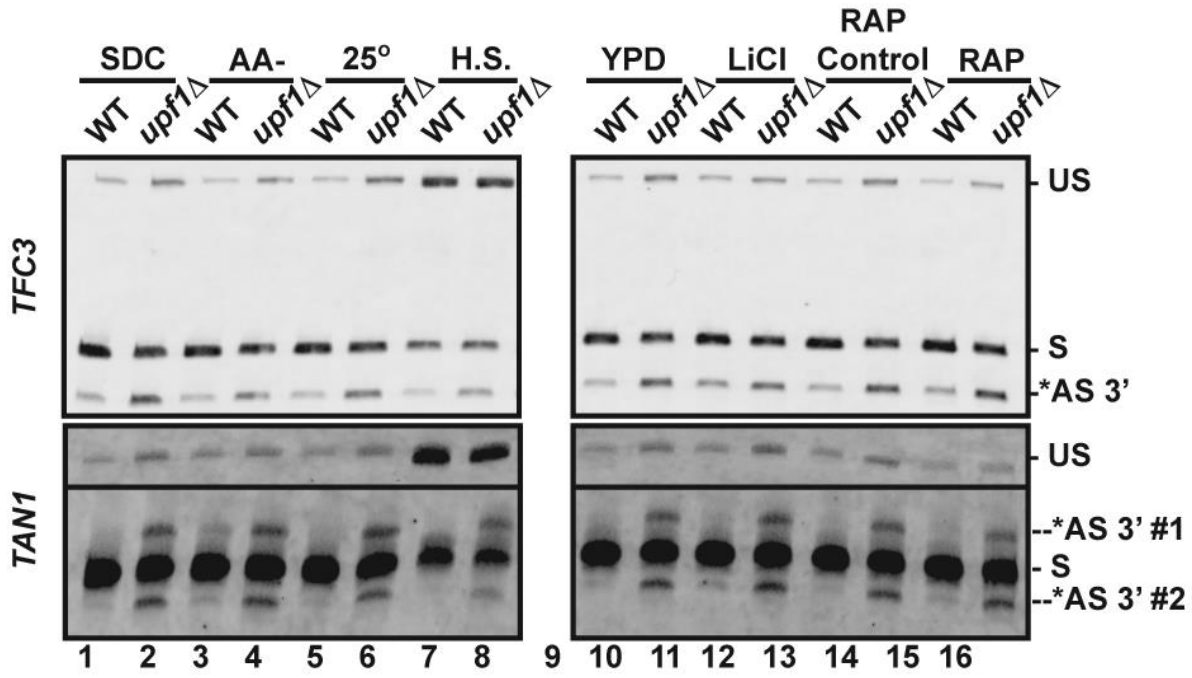

SDC
AA-
25°
H.S.
YPD
LiCl
RAP Control
RAP
WT
upf1Δ
TFC3
TAN1
US
S
*AS 3'
US
--*AS 3' #1
S
--*AS 3' #2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure S12

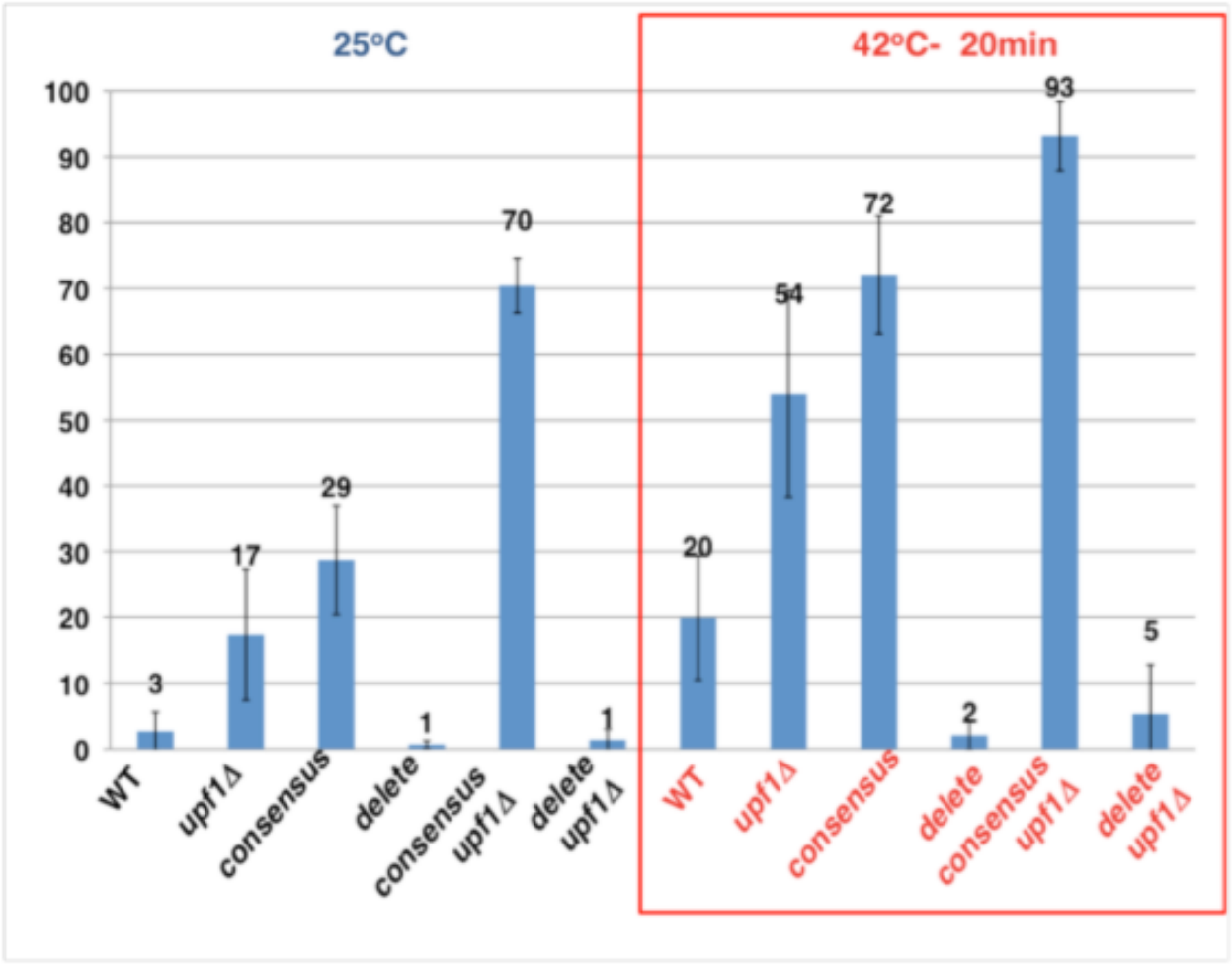
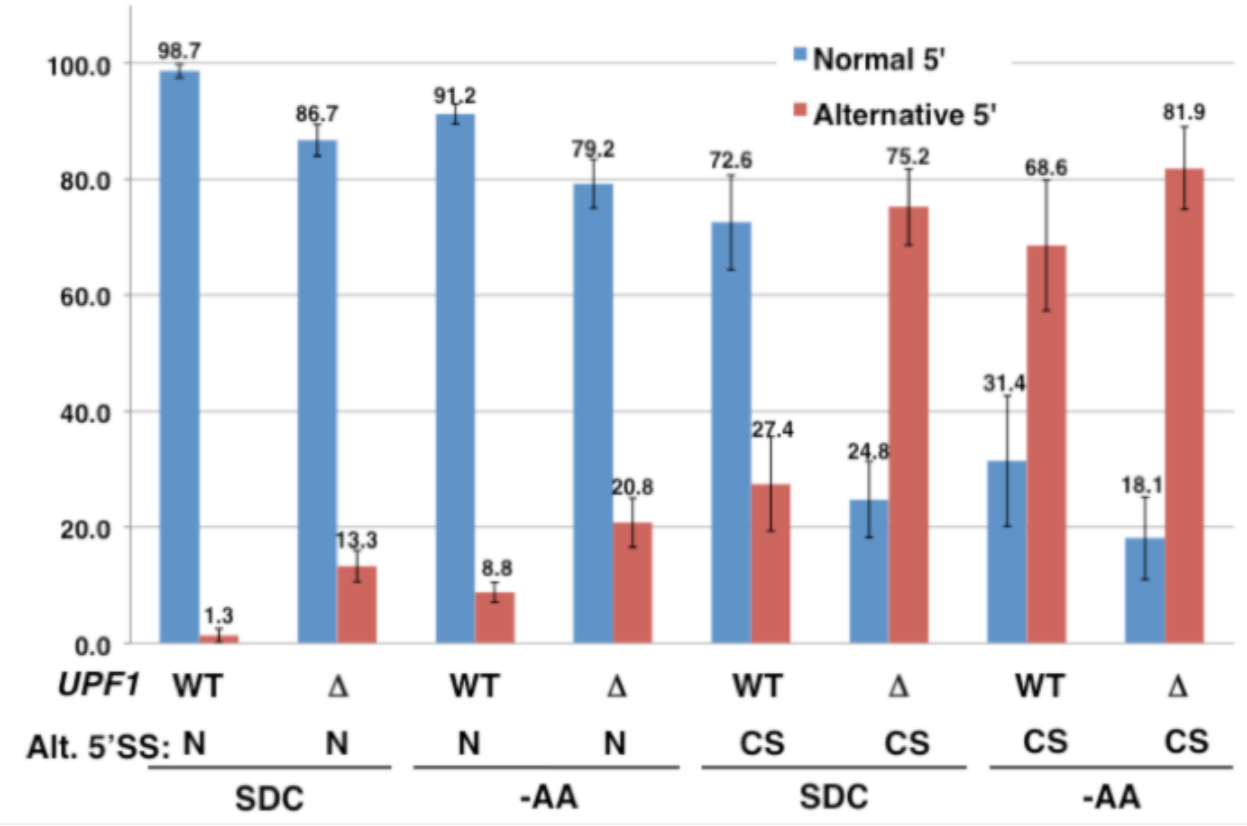

25°C
42°C- 20min
100
90
80
70
60
50
40
30
20
10
0
3
17
29
1
70
1
20
54
72
2
93
5
WT
upf1Δ
consensus
delete
consensus upf1Δ
delete upf1Δ
WT
upf1Δ
consensus
delete
consensus upf1Δ
delete upf1Δ

Figure S13


Normal 5'
Alternative 5'
100.0
80.0
60.0
40.0
20.0
0.0
98.7
1.3
86.7
13.3
91.2
8.8
79.2
20.8
72.6
27.4
24.8
75.2
31.4
68.6
18.1
81.9
UPF1 WT Δ WT Δ WT Δ WT Δ
Alt. 5'SS: N N N N CS CS CS CS
SDC -AA SDC -AA

Chapter 3

Splicing-Mediated Autoregulation Modulates Rpl22p Expression in *Saccharomyces cerevisiae*

PLOS | GENETICS

RESEARCH ARTICLE

Splicing-Mediated Autoregulation Modulates Rpl22p Expression in *Saccharomyces cerevisiae*

Jason Gabunilas[1], **Guillaume Chanfreau**[1,2]*

1 Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California, United States of America, **2** Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California, United States of America

* guillom@chem.ucla.edu

OPEN ACCESS

Citation: Gabunilas J, Chanfreau G (2016) Splicing-Mediated Autoregulation Modulates Rpl22p Expression in *Saccharomyces cerevisiae*. PLoS Genet 12(4): e1005999. doi:10.1371/journal.pgen.1005999

Editor: Hiten D. Madhani, University of California San Francisco, UNITED STATES

Received: October 7, 2015

Accepted: March 30, 2016

Published: April 20, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by grants from the National Institute of General Medical Sciences: grant GM61518 to GC and grant T32 GM007185 to JG. http://www.nigms.nih.gov/Pages/default.aspx. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

In *Saccharomyces cerevisiae*, splicing is critical for expression of ribosomal protein genes (RPGs), which are among the most highly expressed genes and are tightly regulated according to growth and environmental conditions. However, knowledge of the precise mechanisms by which RPG pre-mRNA splicing is regulated on a gene-by-gene basis is lacking. Here we show that Rpl22p has an extraribosomal role in the inhibition of splicing of the *RPL22B* pre-mRNA transcript. A stem loop secondary structure within the intron is necessary for pre-mRNA binding by Rpl22p *in vivo* and splicing inhibition *in vivo* and *in vitro* and can rescue splicing inhibition *in vitro* when added in *trans* to splicing reactions. Splicing inhibition by Rpl22p may be partly attributed to the reduction of co-transcriptional U1 snRNP recruitment to the pre-mRNA at the *RPL22B* locus. We further demonstrate that the inhibition of *RPL22B* pre-mRNA splicing contributes to the down-regulation of mature transcript during specific stress conditions, and provide evidence hinting at a regulatory role for this mechanism in conditions of suppressed ribosome biogenesis. These results demonstrate an autoregulatory mechanism that fine-tunes the expression of the Rpl22 protein and by extension Rpl22p paralog composition according to the cellular demands for ribosome biogenesis.

Author Summary

Ribosomes are responsible for protein production in all living cells, serving as the grounds for the translation of genetic information from RNA to protein. Given the vital role of the ribosome in protein assembly, ribosome components are highly expressed and are subject to tight regulation. Some ribosomal proteins are also known to engage in extra-ribosomal activities. In our study, we demonstrate that the ribosomal protein Rpl22p is able to regulate its own expression by inhibiting the processing of its own RNA transcript, leading to degradation of the RNA. We also show that this self-imposed regulation plays a role in limiting *RPL22* transcript levels in specific stress conditions. We suggest that this

mechanism may impact the composition of ribosomes by influencing the availability of the Rpl22p paralogs.

Introduction

Ribosomal protein genes (RPGs) constitute a majority of the most frequently transcribed genes in the budding yeast *Saccharomyces cerevisiae* [1]. Due in part to their high levels of expression and their vital role as components of the translational machinery, understanding of the regulation of RPG expression has garnered considerable attention. While RPGs are tightly regulated at the transcriptional level [2], the fact that nearly half of all intron-containing genes in *S. cerevisiae* are RPGs [3] has led to questions regarding the importance of these introns in RPG regulation. To address this, a comprehensive deletion of the yeast RPG intronome revealed numerous cases of intron-dependent intergenic and intragenic regulation of RPG expression that also impacted cell growth in various stress conditions [4]. These findings led to the conclusion that introns within RPGs govern the auto- and cross-regulation of RPG expression. While some structural elements within intronic RPGs were found to be important for splicing efficiency [5,6], the precise mechanisms by which this regulation is achieved on a gene-by-gene basis remain largely unknown.

Over the past several decades, a few studies have shown that the regulation of expression of particular RPGs is in part dependent upon extra-ribosomal autoregulatory functions of the ribosomal proteins themselves [7]. In *S. cerevisiae*, some ribosomal proteins have been found to regulate the splicing of their own pre-mRNA transcripts. Such direct regulation has been characterized for the ribosomal protein Rpl30p, which binds to an RNA secondary structure in the precursor transcript and inhibits its splicing [8]. Similarly, the ribosomal protein Rps14p self-regulates its expression by directly binding to a stem loop structure in the *RPS14B* pre-mRNA [9], while the ribosomal protein Rps9p preferentially represses splicing of the *RPS9A* minor paralog through the recognition of an intronic structural element [10]. Other ribosomal proteins have been found to regulate their mRNAs by mechanisms other than splicing, particularly in cases where the nascent transcript does not contain an annotated intron. For instance, recent studies showed that yeast Rps28p indirectly binds a regulatory element in the 3' untranslated region (3'UTR) of its mRNA transcript via Edc3p and targets the mRNA for decapping and degradation [11], while Rpl9p influences the transcription termination pathway of the *RPL9B* transcript, coupling termination to nuclear degradation [12]. RPG autoregulation is not limited to *S.cerevisiae* but has also been identified in higher eukaryotes. In mice and zebrafish the ribosomal protein Rpl22 regulates the expression of its paralog protein Rpl22l1 by interacting with the Rpl22l1 pre-mRNA, thereby repressing expression of the protein through an as-yet unknown mechanism [13].

We previously showed that the pre-mRNA of *RPL22B* contains an intronic alternative 5' splice site and that splicing at this site gives rise to a transcript that is degraded by the cytoplasmic nonsense-mediated decay (NMD) pathway [14]. This finding suggested that alternative splicing of this precursor transcript may serve as a means for regulating mature *RPL22* transcript levels in an NMD-dependent manner. In this study, we describe an autoregulatory circuit for the regulation of *RPL22* in *S.cerevisiae* based on the inhibition of the splicing of the *RPL22B* pre-mRNA by Rpl22p. We identify and characterize an RNA stem loop within the *RPL22B* intron that is necessary for the inhibition of pre-mRNA splicing by Rpl22p *in vivo* and *in vitro*. Physiologically, we found that this mechanism promotes the down-regulation of spliced *RPL22B* transcript during stress. Together with our previous findings, these results

demonstrate that *RPL22B* is precisely regulated at the RNA level by multiple splicing-based mechanisms and identify a physiological extraribosomal function of Rpl22p during stress.

Results

Splicing of the *RPL22B* pre-mRNA is regulated by the Rpl22p protein

For several duplicated genes in *S. cerevisiae*, including *RPL22*, deletion of one paralog results in a compensatory enhancement of expression of the remaining paralog [15,16,17]. However, the mechanisms by which this upregulation occurs are not always clear. Because the vast majority (>90%) of ribosomes contain Rpl22p produced from the *RPL22A* locus in wild-type yeast cells [18], we hypothesized that the loss of this gene may trigger a compensatory response with regards to expression and/or processing of the paralogous *RPL22B* transcript.

To determine whether Rpl22p-mediated splicing regulation occurs for *RPL22B*, we performed RT-PCR using locus-specific *RPL22B* primers on cDNA derived from total RNA of wild-type cells and cells harboring the deletion of the *RPL22A* gene (*rpl22aΔ*). The deletion of *RPL22A* resulted in a substantial increase in the proportion of spliced transcript relative to the unspliced pre-mRNA and the 5' alternatively spliced *RPL22B* transcript that is subject to NMD (Fig 1A). Because of the severe growth defect of the *rpl22aΔ* strain [17], we could not rule out that this splicing pattern was an indirect effect of slow growth, as growth impairment has been shown to confer multiple phenotypes in yeast including resistance to heat shock and ER stress [17,19]. Therefore, we also analyzed *RPL22B* splicing patterns in two other slow-growing RP mutants, *rpl31aΔ* and *rpl39Δ* ([20], kindly provided by B. Kennedy). We found that the splicing pattern for *RPL22B* was remarkably different in these two mutants in comparison to wildtype and *rpl22aΔ*. Specifically, these mutants both displayed an accumulation of the unspliced and alternative spliced transcripts while the proportion of spliced transcript was concomitantly reduced (Fig 1A). Therefore, the splicing pattern for *RPL22B* in the *rpl22aΔ* strain appears to be unique to that strain and does not result from a severe growth defect. To ensure that the reduction in pre-mRNA levels detected in the *rpl22aΔ* strain was specific to *RPL22B*, we repeated the RT-PCR analyses on each of these strains for *TFC3*, which contains an alternative 3' splice site [14]. We found no observable differences in *TFC3* splicing patterns across these four strains (Fig 1B) in spite of considerable differences in generation time [20]. This result suggests that the splicing pattern for *RPL22B* in the *rpl22aΔ* mutant is unique to that strain and that the deletion of *RPL22A* does not confer a general change in splicing patterns. Interestingly, the levels of the 5' alternatively spliced species did not increase in the *rpl22aΔ* strain but was instead decreased (Fig 1A, lane 2), suggesting that the inhibition of splicing is specific to the annotated splice site.

The precursor and alternatively spliced transcripts of *RPL22B* are degraded by NMD [14,21], making their accumulation difficult to detect in the context of an active NMD system. To better characterize the impact of *RPL22A* deletion on the splicing of *RPL22B*, we analyzed *RPL22B* levels by Northern blot in both wild-type and NMD-deficient (*upf1Δ*) strains in the presence or absence of *RPL22A*. As expected, disabling NMD led to the accumulation of the NMD-sensitive unspliced transcript while the deletion of *RPL22A* reduced the unspliced signal to undetectable levels (Fig 1C), in accordance with the results from the RT-PCR. Furthermore, the deletion of *RPL22A* attenuated the accumulation of the unspliced transcript in the *upf1Δ* mutant (Fig 1C, compare lanes 2 and 4), either because the degradation of this transcript is hyperactivated in the *rpl22aΔ* mutant or, alternatively, because the absence of *RPL22A* allows for more efficient splicing of the pre-mRNA. The latter hypothesis was supported by the observation that the *rpl22aΔ* deletion also resulted in a significant increase in the levels of spliced *RPL22B* transcript compared to wildtype (Fig 1C, lanes 1 and 3), a difference that was not

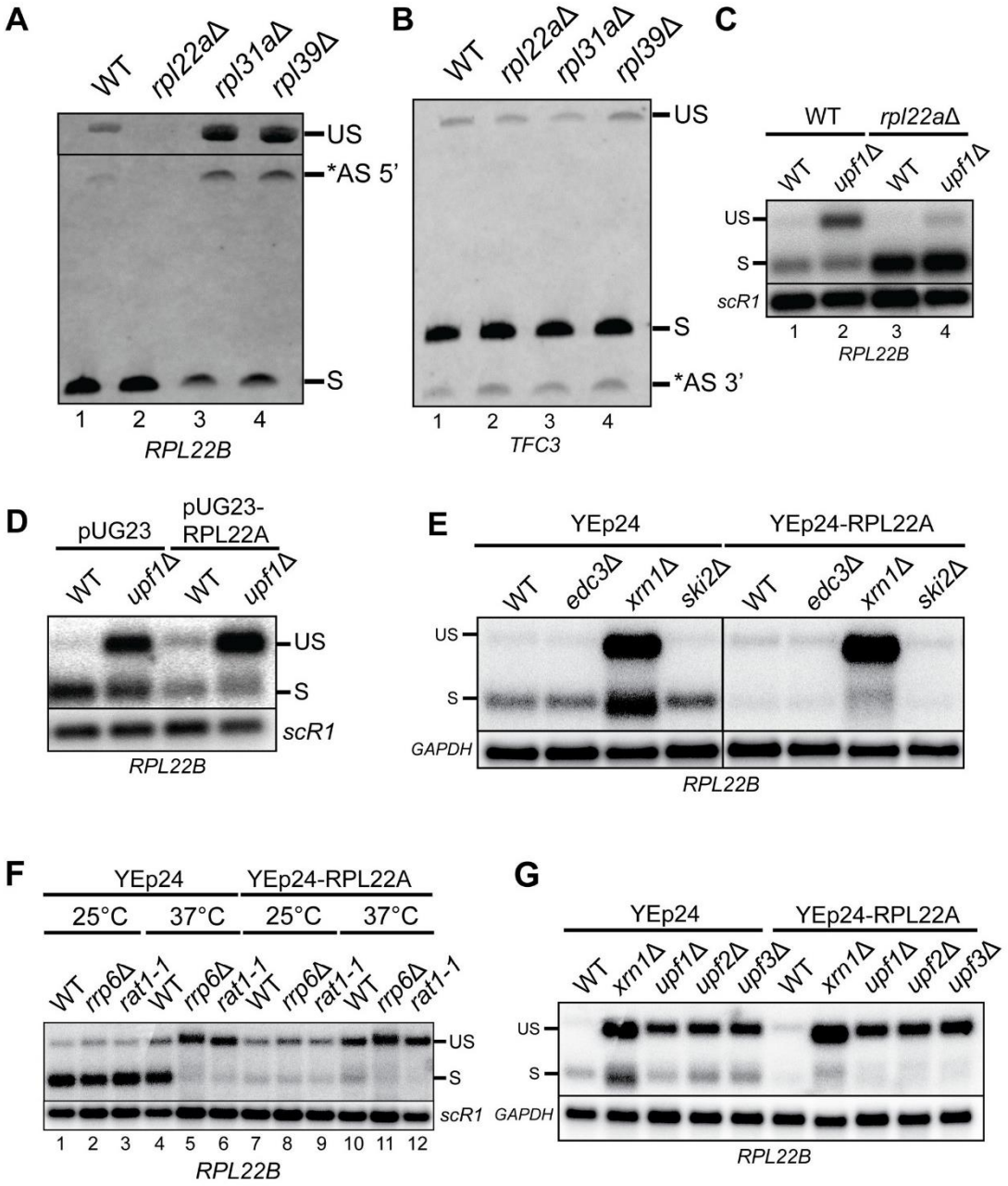


Fig 1. Rpl22Ap modulates the splicing of *RPL22B* pre-mRNA, which is targeted for degradation by cytoplasmic and nuclear decay factors. A. RT-PCR analysis of splicing products of the *RPL22B* pre-mRNA in wild-type and ribosomal protein deletion mutants. Bands indicate the unspliced (US), alternatively-spliced (*AS 5'), and spliced (S) species. **B.** RT-PCR analysis of splicing products of the *TFC3* pre-mRNA in wild-type and ribosomal protein deletion mutants. Bands indicate the unspliced (US), spliced (S), and alternatively spliced (*AS 3') species. **C.** Northern blot analysis of *RPL22B* splicing products in wild-type, *upf1*Δ, *rpl22a*Δ, and *upf1*Δ/*rpl22a*Δ cells detected using an *RPL22B* 5'UTR riboprobe. Shown are the unspliced (US) and spliced (S) species. *SCR1* was used as a loading control. **D.** Northern blot showing *RPL22B* splicing products in wild-type or *upf1*Δ cells with an empty pUG23 vector or intronless *RPL22A* cDNA overexpression plasmid. Splicing products were detected using an *RPL22B* 5'UTR riboprobe. Labeling of the transcripts is similar

to panel C. *SCR1* was used as a loading control. **E.** Northern blot analysis of *RPL22B* in cytoplasmic RNA decay mutants carrying the multi-copy YEp24 vector with either no insert or expressing the intronless *RPL22A* gene. Labeling of the transcripts is similar to panels C-D. *GAPDH* was used as a loading control. **F.** Northern blot analysis of *RPL22B* in nuclear RNA decay mutants carrying an empty YEp24 vector or an intronless *RPL22A* overexpression plasmid. Strains were grown to exponential phase in–URA media at 25°C and then shifted to 37°C for three hours. Labeling of the transcripts is similar to panels C-E. *SCR1* was used as a loading control. **G.** Northern blot analysis of *RPL22B* in *xrn1*Δ and NMD mutant strains carrying an empty YEp24 vector or an *RPL22A* overexpression plasmid. Detection methods and labeling are similar to panels C-F. *GAPDH* was used as a loading control.

doi:10.1371/journal.pgen.1005999.g001

detected by RT-PCR in Fig 1A, likely because of saturated amplification of the more abundant spliced cDNA in both samples.

Because the deletion of *RPL22A* enhanced the levels of spliced *RPL22B* transcript and repressed the levels of unspliced transcript, we reasoned that overexpressing Rpl22p would have the opposite effect and repress the levels of spliced transcript. To test this hypothesis, we created an *RPL22A* cDNA overexpression vector in which an intronless *RPL22A* cDNA was cloned into the pUG23 plasmid and placed under the control of the *MET25* promoter such that yeast cells transformed with this plasmid overexpress the *RPL22A* cDNA when grown in media lacking methionine [22]. Indeed, overexpression of the *RPL22A* cDNA resulted in the repression of levels of spliced *RPL22B* transcript and accumulation of the pre-mRNA (Fig 1D lanes 1 and 3). Notably, the accumulation of the pre-mRNA detected in the *upf1*Δ strain was even more pronounced when *RPL22A* was overexpressed (Fig 1D, lanes 2 and 4). We note that the impact of *RPL22A* overexpression on the *RPL22B* transcript splicing was not as dramatic as the effect of the *RPL22A* deletion, likely because the overexpression of an already highly-expressed gene such as *RPL22A* induces a minor perturbation compared to the deletion of the gene. Altogether, these results suggest that the splicing efficiency of the *RPL22B* transcript is heavily influenced by the levels of Rpl22p protein, the majority of which is derived from the *RPL22A* locus.

Nuclear and cytoplasmic decay of the *RPL22B* pre-mRNA is not influenced by Rpl22p autoregulation

To determine whether other mechanisms contribute to the degradation of the unspliced *RPL22B* transcript, we analyzed the accumulation of the *RPL22B* pre-mRNA in deletion mutants for the cytoplasmic RNA decay factors Xrn1p [23] and the Ski2p activator of the exosome [24], as well as the mRNA decapping factor Edc3p which is involved in autoregulation of *RPS28B* [11]. We analyzed *RPL22B* pre-mRNA levels in these deletion mutants with or without the overexpression of intronless *RPL22A* from a multicopy YEp24 plasmid to determine whether the inhibition of splicing impacts the route of degradation taken by the pre-mRNA.

As expected, inactivation of Xrn1p resulted in an increase of both unspliced and spliced *RPL22B* transcript (Fig 1E), consistent with the general role of Xrn1p in cytoplasmic mRNA degradation [25]. By contrast, the absence of Edc3p or Ski2p had no effect on *RPL22B* pre-mRNA levels, suggesting that this transcript is not a major target of Edc3-mediated decapping or of the cytoplasmic exosome. Overexpression of *RPL22A* within this context repressed spliced *RPL22B* levels in all strains but did not affect the sensitivity of the pre-mRNA to any of the decay factors (Fig 1E). Therefore, splicing inhibition by Rpl22p overexpression does not appear to alter the decay fate of cytoplasmic *RPL22B* pre-mRNA.

To ascertain the sensitivity of *RPL22B* pre-mRNA to nuclear decay pathways, we analyzed the impact of the nuclear exosome component deletion strain *rrp6*Δ and of the temperature-sensitive 5' to 3' nuclear exonuclease mutant *rat1-1*. No differences in *RPL22B* levels were detected between the wild-type strain and these two mutants at the permissive temperature of 25°C (Fig 1F, lanes 1–3). Shifting wildtype and *rrp6*Δ strains to 37°C for three hours resulted in an observable increase in the levels of pre-mRNA (Fig 1F, lanes 4–5) which was even more

pronounced in the *rrp6*Δ strain, consistent with our previous studies [26]. Likewise, shifting *rat1-1* to the non-permissive temperature also resulted in a substantial accumulation of unspliced transcript that also exceeded that seen in wildtype (Fig 1F, lane 6). While these changes at 37°C could be partially attributed to heat shock-induced inhibition of splicing [27,28], the higher levels of unspliced transcript present in the decay mutants in comparison to wildtype suggest that this pre-mRNA is targeted by both 5' to 3' and 3' to 5' nuclear exonucleolytic degradation mechanisms in heat shock conditions. Interestingly, the levels of spliced *RPL22B* were dramatically reduced in *rrp6*Δ and *rat1-1* at 37°C, whereas the wild-type strain maintained high levels of that transcript (Fig 1F, lanes 4–6). The reason for this effect is unclear, but it is possible that the wild-type strain is able to recover from the initial heat shock and re-establish normal splicing patterns more rapidly than either of the mutant strains. Alternatively, the levels of Rpl22p may be enhanced in nuclear RNA decay factor mutants. Importantly, overexpressing *RPL22A* reduced the levels of spliced transcripts across all strains at both temperatures (Fig 1F, lanes 7–12), but did not appear to influence the susceptibility of the pre-mRNA to degradation by either exonuclease in particular.

In higher eukaryotes the NMD helicase Upf1p is involved in multiple nuclear processes including DNA replication and repair and cell cycle control in addition to several NMD-independent RNA decay pathways (see [29,30] for review). Given our finding that the *RPL22B* pre-mRNA is subject to nuclear RNA degradation, we sought to determine whether the accumulation of the unspliced transcript previously detected in the *upf1*Δ strain is indeed a direct consequence of the lack of NMD activity and not the result of NMD-independent roles of Upf1p. We analyzed three NMD-compromised strains, each lacking one of the three up-frameshift proteins, *upf1*Δ, *upf2*Δ, and *upf3*Δ, along with *xrn1*Δ. All three NMD mutants exhibited a very similar accumulation of the pre-mRNA (Fig 1G), confirming that degradation of the unspliced *RPL22B* pre-mRNA occurs through NMD. Notably, the genetic inactivation of Xrn1p shows a stronger accumulation of pre-mRNA relative to the upstream NMD factors in addition to the accumulation of spliced transcript. When the degradation rate of an RNA transcript is altered, an Xrn1p-dependent RNA buffering mechanism modulates the rate of the RNA synthesis in a compensatory manner [25,31]. Therefore, lack of RNA buffering in the *xrn1*Δ strain to may contribute to the stronger accumulation of these transcripts in that strain relative to the NMD mutants.

Altogether, our results demonstrate that both nuclear and cytoplasmic RNA decay mechanisms are responsible for the degradation of *RPL22B* pre-mRNA, though it is more heavily targeted by the latter. Furthermore, the degradation pathway taken by the transcript does not appear to be influenced by Rpl22p levels.

Splicing of *RPL22A* pre-mRNA is mildly affected by Rpl22p

The *RPL22* paralogs share a high percentage of nucleotide identity within their protein coding sequences and a high percentage of amino acid identity (S1A Fig). We therefore wondered whether the protein product of *RPL22B* is capable of inhibiting the splicing of the *RPL22B* pre-mRNA. To test this, we exogenously overexpressed intronless *RPL22B* cDNA from a plasmid in either wild-type or NMD mutant strains and analyzed *RPL22B* mRNA splicing patterns by Northern blot. Strikingly, the overexpression of *RPL22B* strongly inhibited the splicing of *RPL22B* pre-mRNA (S1B Fig), demonstrating that the inhibitory property is shared between both protein paralogs. We also tested whether the autoregulation affects the *RPL22A* pre-mRNA by analyzing the splicing of *RPL22A* pre-mRNA upon chromosomal deletion or exogenous overexpression of *RPL22B* or exogenous overexpression of intronless *RPL22A* (S1C–S1E Fig). These experiments revealed a minor repression of spliced *RPL22A* mRNA and slight increase in pre-mRNA levels upon overexpression of either *RPL22* paralog (S1C and S1D Fig),

but no affect upon genetic inactivation of *RPL22B* (S1E Fig). Therefore, the *RPL22A* pre-mRNA appears to be mildly susceptible to the autoregulatory activity, though not nearly to the extent seen in *RPL22B*.

Regulation of *RPL22B* splicing by Rpl22p is mediated by the *RPL22B* intron

Because introns within many RPGs are capable of regulating both inter- and intragenic expression [4], we asked whether the intronic sequence of *RPL22B* is necessary and sufficient for splicing regulation by Rpl22p. To answer this question, we created a chromosomally-expressed chimeric gene at the *RPL18B* locus using the *delitto perfetto* technique [32] in which we replaced the intron of *RPL18B* with the intron from *RPL22B* (Fig 2A). In addition to wild-type strains, these mutations were also made in the context of NMD deficiency, deletion of *RPL22A*, or both (respectively, *upf1*Δ, *rpl22a*Δ, and *upf1*Δ*rpl22a*Δ). As expected, the deletion of *RPL22A* did not increase the splicing efficiency for wild-type *RPL18B* pre-mRNAs, and in fact, appeared to slightly impair splicing (Fig 2B, compare lanes 1 and 2 with 3 and 4). We speculate that this may be an indirect effect of slower growth, as evidenced by the remarkably different splicing patterns for *RPL22B* observed in the slow-growing strains *rpl31a*Δ and *rpl39*Δ compared to wildtype (Fig 1A). More importantly, substitution of the *RPL22B* intron into *RPL18B* reduced the amount of spliced transcript generated from the chimeric pre-mRNA, likely due to splicing inhibition by Rpl22p (Fig 2B, compare lanes 1 and 2 with lanes 5 and 6). In support of this notion, the deletion of *RPL22A* in the context of the chimeric *RPL18B*::*RPL22B* intron gene resulted in a substantial increase in the amount of spliced *RPL18B* transcript (Fig 2B, lanes 7–8), similar to what was observed for the native *RPL22B* transcript (Fig 1C).

To further assess that Rpl22p-mediated regulation can operate *in trans*, we performed the complementary experiment in which *RPL22A* was overexpressed in either wild-type or *upf1*Δ strains expressing native or chimeric *RPL18B*. Rpl22p overexpression had no discernible impact on the splicing of the wild-type *RPL18B* transcript, demonstrating the specificity of this regulatory mechanism (Fig 2C, lanes 1 through 4). By contrast, overexpressing *RPL22A* reduced the levels of spliced chimeric transcript while also increasing the levels of unspliced precursor in the wild-type strain (Fig 2C, compare lanes 5 and 7). Taken together, these experiments suggest that the Rpl22p-mediated splicing regulation of *RPL22B* is mediated by its intron and can operate *in trans* to inhibit the splicing of another transcript harboring the *RPL22B* intron.

A region within the *RPL22B* intron harboring a stem loop structure is necessary for splicing autoregulation by Rpl22p

Previous studies have demonstrated that both direct and indirect autoregulation of genes in yeast are often dependent upon RNA secondary structures [8,11,13]. Based on these precedents, we investigated the presence of a structural regulatory element within the *RPL22B* intron. To simplify the process of manipulating the intron, we first constructed a plasmid-based reporter in which the *RPL22B* intron was attached to a transcript encoding green fluorescent protein (GFP) (Fig 3A). We then utilized site-directed mutagenesis to sequentially bifurcate the intron and assessed loss of splicing inhibition by comparing relative levels of spliced and unspliced reporter transcript by Northern blot (Fig 3A). We excluded the branch point sequence and 5' and 3' splice sites from any of the deletions in order to not disrupt the splicing mechanism. Through this approach, we isolated a region spanned by intronic nucleotides 153 through 225 that harbors a prominent secondary RNA structure (as predicted by Mfold [33]) and that was necessary to confer splicing autoregulation. This structure consists of a series of

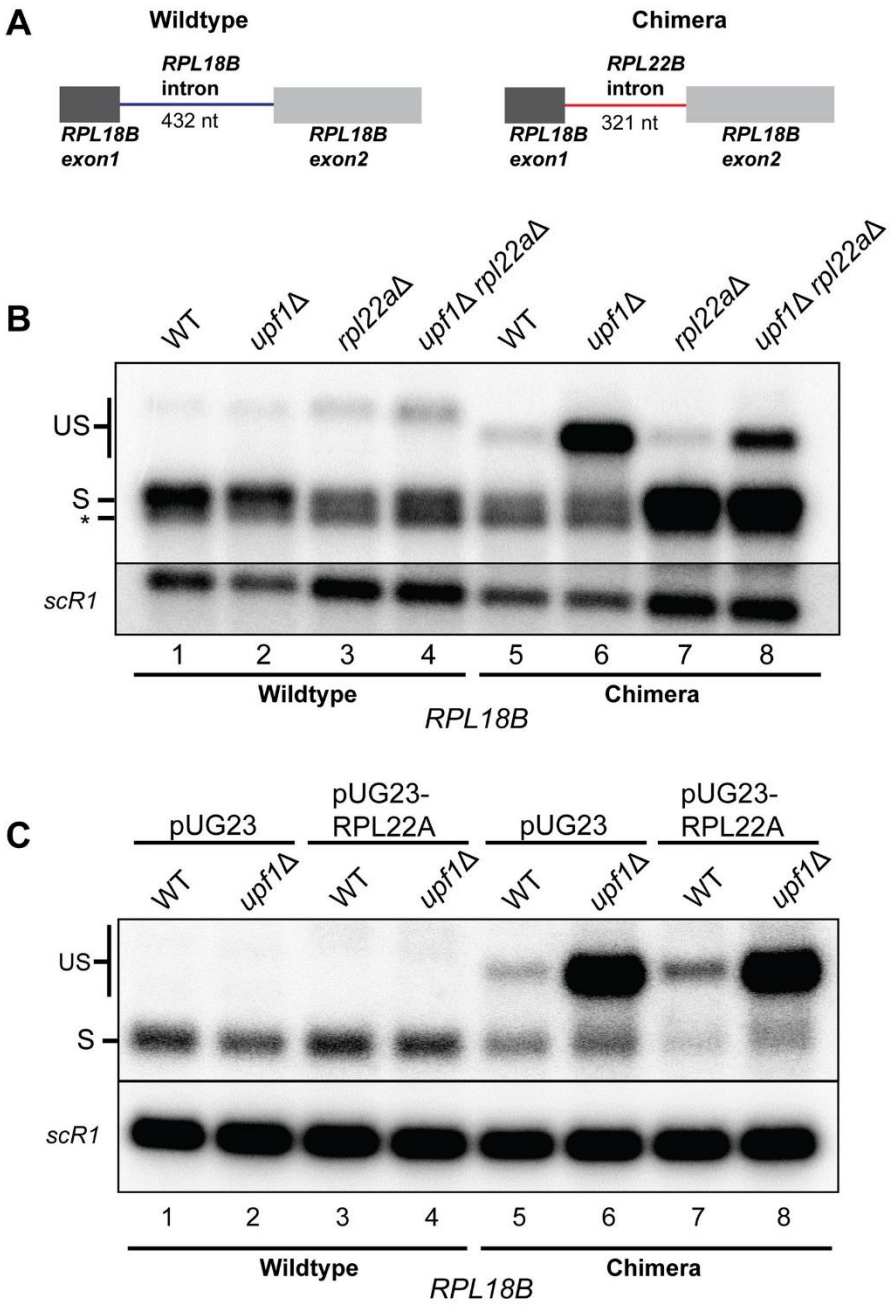


Fig 2. Rpl22p-mediated regulation of splicing is intron-specific. A. Schematic of chromosomal wild-type and chimeric *RPL18B* genes. In the chimera, the intron from *RPL18B* was replaced with the intron from *RPL22B* using the *delitto perfetto* approach. **B.** Northern blot detecting the splicing patterns for wild-type and chimeric *RPL22B* transcripts in wild-type and NMD mutant strains with or without the deletion of *RPL22A*. Labeled bands represent the wild-type and chimeric unspliced transcripts (US) and the spliced transcript (S). Transcripts were detected using an *RPL18B* riboprobe that spans the 3' end of the ORF and the 3'UTR. The asterisk indicates non-specific detection of the *RPL18A* spliced transcript by the *RPL18B* probe. *SCR1* was used as a loading control. **C.** Northern blot detecting the splicing patterns for wild-type and chimeric *RPL22B* transcripts in WT and NMD mutants with or without the overexpression of intronless *RPL22A* from pUG23. Labeling and detection of bands is similar to panel B. *SCR1* was used as a loading control.

doi:10.1371/journal.pgen.1005999.g002

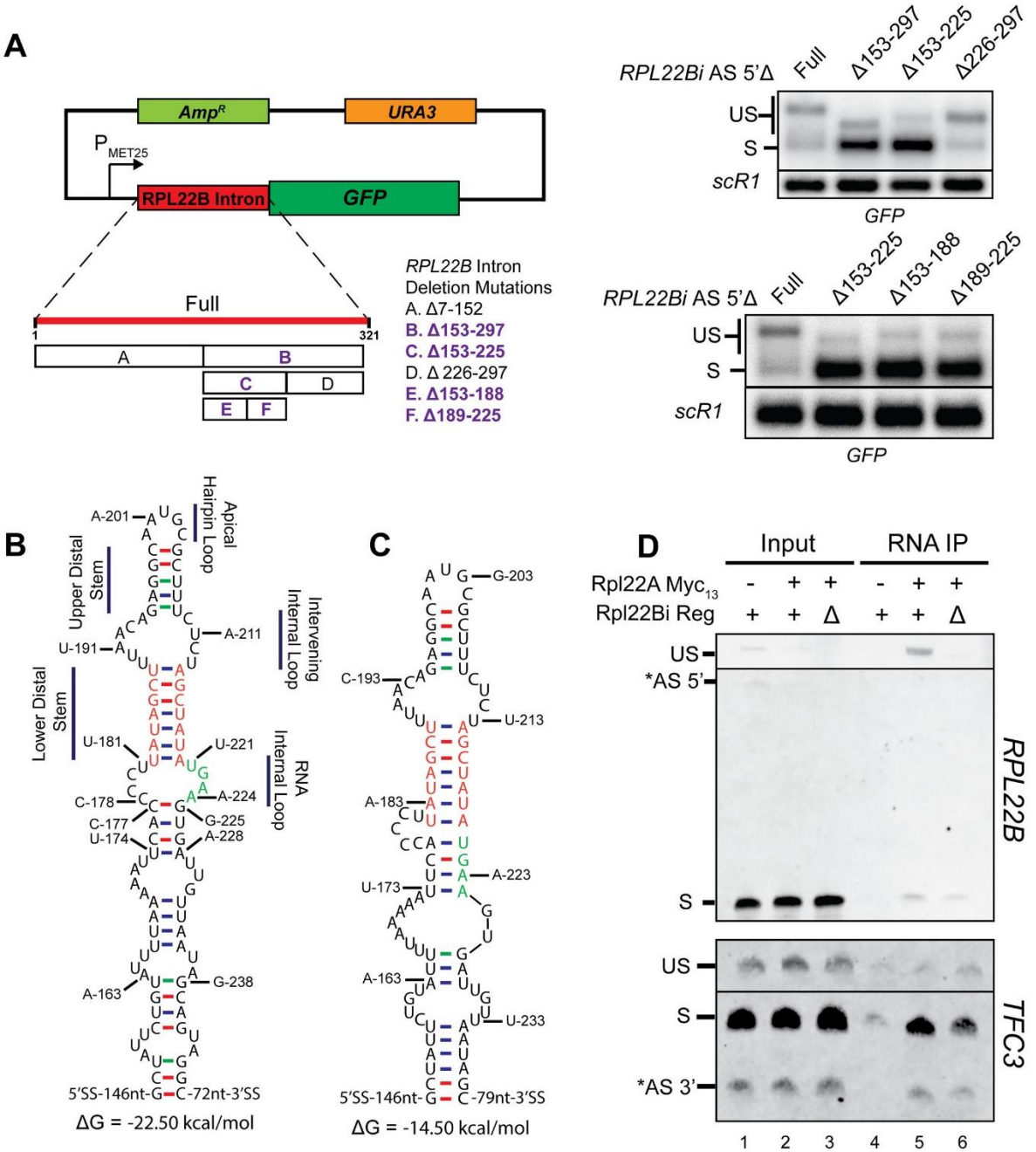


Fig 3. Identification of a regulatory element necessary for binding of Rpl22 to the unspliced *RPL22B* mRNP *in vivo*. A. *Left*: Schematic of *RPL22B* intron reporter system based on a pUG35 plasmid backbone. Also indicated are the deletions of the binary search. Deletions highlighted in bold purple font indicate those that result in a loss of splicing inhibition. *Right*: Northern blots demonstrating the loss of splicing inhibition upon deletion of certain segments of the *RPL22B* intron. *SCR1* was used as a loading control. Note that all constructs appearing in these blots do not contain the alternative 5'-splice site. See S1 Text for details. **B.** Structure of the *RPL22B* intron bounded by nucleotides G-153 and C-246 as predicted by Mfold. Regions of interest have been denoted.

Nucleotides of the lower distal stem and nucleotides U221 through A224 of the putative RNA Internal Loop were experimentally determined to be important for autoregulatory activity and are colored orange and green, respectively. **C.** Experimentally deduced intronic regulatory element that mediates the inhibition of splicing of the *RPL22B* pre-mRNA. Nucleotide colors from Panel B are retained for comparison. Details describing the experiments leading to the inferred secondary structure are provided in S1 Text. **D.** RNA immunoprecipitation shows the association of Rpl22p with the unspliced *RPL22B* mRNP that requires the regulatory element. *Upper panel*: RT-PCR analysis of *RPL22B* transcripts derived from cells expressing untagged Rpl22Ap (-), Rpl22Ap with the 13-Myc epitope tag (+), and cells with the *RPL22B* intronic regulatory element deleted (Δ) and expressing tagged protein (+). Total input RNA and immunoprecipitated RNA were analyzed from each strain. Labeled bands show the unspliced (US) and spliced (S) species. *Lower panel*: RT-PCR analysis of *TFC3* transcripts performed by similar methods as described for panel A. Bands indicate the unspliced (US) and spliced (S) species. The gel images have been truncated for space considerations. Full, unattenuated versions of the gels can be viewed in S8B Fig.

doi:10.1371/journal.pgen.1005999.g003

base-paired stems and internal loops closed by an apical hairpin loop (Fig 3B). A more comprehensive description of the binary search approach which led to the identification of this element can be found in S1 Text and S2 Fig.

While this approach allowed us to narrow the window in which the regulatory element resides, it did not provide the necessary resolution to precisely determine the necessary features that enable autoregulation. Therefore, we opted to continue with specific targeted mutations of the putative structure isolated by the binary search. In brief, we found that the identities of the component nucleotides comprising what we refer to as the lower distal stem are inconsequential in governing *RPL22B* regulation, as long as the stem itself is able to form (Fig 3B and 3C [stem nucleotides in red], S4E and S4F Fig). We also determined that efficient autoregulation hinges on the identities of the downstream nucleotides comprising a predicted RNA internal loop, which we later determined experimentally to be paired with upstream nucleotides to form a short stem (highlighted in green in Fig 3B and 3C; see also S5 Fig and S1 Text). The experimentally-deduced structure of the regulatory element is presented in Fig 3C and the details of these experiments are described extensively in S1 Text. We could not demonstrate that this sequence was sufficient to confer splicing inhibition *in vivo* when transposed into a different intron (see S6 Fig, S1 Text and Discussion), possibly because additional intronic elements are required or because the stem-loop is unable to fold correctly in the context of a different transcript. However, a short RNA sequence containing the regulatory element was able to enhance splicing efficiency *in vitro* when added *in trans* to splicing reactions, suggesting that this element was sufficient to titrate Rpl22p in extracts, and thus corresponds to a bona fide binding element for Rpl22p (see below).

Rpl22p associates with unspliced *RPL22B* mRNPs via the intronic regulatory element

Our previous observations support a splicing-mediated autoregulatory mechanism for *RPL22B*, wherein the changes in splicing efficiency of the *RPL22B* pre-mRNA result from a direct or indirect physical association between the pre-mRNA and protein. To definitively demonstrate this association *in vivo*, we immunoprecipitated the endogenous Rpl22p protein produced from the *RPL22A* locus, using a 13-Myc epitope tag fused to the C-terminus of the Rpl22Ap protein (S8A Fig), followed by RNA analysis (RNA immunoprecipitation or RIP). The RIP was performed on a strain expressing untagged protein as a negative control, the Rpl22Ap-Myc tagged strain expressing wild-type *RPL22B* pre-mRNA, and an Rpl22A-Myc tagged strain expressing the *RPL22B reg*Δ pre-mRNA as an additional negative control. RNA pulldown was followed by RT-PCR analysis.

As expected, analysis of the total input RNA from all three strains indicated that the mature spliced species is the most abundant form of the *RPL22B* transcript in the cell (Fig 3D, upper panel, lanes 1–3). Immunoprecipitation of RNA from the strain expressing untagged protein did not yield any detectable PCR products. However, pulldown of Myc-tagged Rpl22Ap revealed a

substantial enrichment of the unspliced *RPL22B* (Fig 3D, lanes 4–5). Because this approach also indirectly precipitates intact ribosomes and any associated translating mRNAs, this enrichment is remarkable in that it clearly exceeds the indirect (ribosome-mediated) association of Rpl22p with the more abundant spliced species that are in the process of being translated. Importantly, the removal of the regulatory element from the *RPL22B* intron eliminated this enrichment for the pre-mRNA in the pulldown (Fig 3D, lane 6), confirming that the regulatory element is required for *in vivo* association between Rpl22p and the unspliced *RPL22B* mRNPs. To further ensure that this interaction is specific for the *RPL22B* pre-mRNA, we also analyzed Rpl22p association with *TFC3* by RT-PCR. As anticipated, pulldown of tagged Rpl22p mostly detected a ribosome-mediated interaction with the canonically spliced *TFC3* transcript, a preference that was unchanged by the removal of the *RPL22B* intronic regulatory element (Fig 3D, bottom panel). We conclude that the *RPL22B* intronic regulatory element is necessary to mediate an interaction between the Rpl22 protein and the unspliced *RPL22B* mRNP *in vivo*.

We next asked whether a structure similar to the identified regulatory element may be governing the interaction between Rpl22p and the large ribosomal RNA (rRNA) in the intact ribosome, as this would represent a logical extension of the protein's natural interactions with rRNA. Examination of published crystallographic and computational structural data [34,35] describing the large ribosomal subunit (LSU) reveals that Rpl22p is an external ribosomal protein that indeed appears to be in contact with a stem loop of the LSU rRNA (S9A Fig). This structure bears no discernible similarity to the regulatory element identified by our studies (Figs S9B and 3C). It does, however, appear to comply with the generalized human Rpl22 RNA-binding motif that was previously identified by SELEX [36] (see Discussion), suggesting that the mRNA binding propensity of mammalian Rpl22 stems from its native interactions with rRNA. The fact that the *RPL22B* regulatory element identified in *S.cerevisiae* does not resemble these sequences or structures may suggest that the RNA binding specificity of the Rpl22p protein has uniquely evolved in the yeast system or that the binding of Rpl22p to the unspliced *RPL22B* mRNP may be indirect and involve a mediating factor. Alternatively, we cannot rule out that the yeast intronic regulatory element might adopt a fold similar to that found in the large rRNA but that this cannot be predicted from its secondary structure.

Regulation of *RPL22B* does not depend on cytoplasmic localization

The apparent "inhibition" of splicing that we detected upon Rpl22 overexpression would also be predicted if Rpl22p expedites the nuclear export of the pre-mRNA and sequesters it out of reach of the spliceosome, a scenario made possible by the trafficking of ribosomal proteins in and out of the nucleus during ribosome maturation [37,38]. To test whether Rpl22p-mediated autoregulation operates when the *RPL22B* unspliced pre-mRNA is constrained in the nucleus, we analyzed the splicing of *RPL22B* pre-mRNA in strains containing a temperature-sensitive mutation of the mRNA export factor *MEX67* [26]. As expected, overexpression of *RPL22A* resulted in the repression of spliced *RPL22B* mRNA in strains expressing either wild-type *MEX67* or the mutant allele *mex67-5* at the permissive temperature (Fig 4A, lanes 1–4). However, the levels of spliced transcript were repressed in the *mex67-5* mutant upon shifting to the non-permissive temperature with or without *RPL22A* overexpression (Fig 4A lanes 7–8 and Fig 4B lanes 5–6), suggesting that this transcript might be subject to turnover by the nuclear exosome in these conditions. Indeed, genetic inactivation of Rrp6p resulted in increased levels of both spliced and unspliced *RPL22B* mRNA (Fig 4B, lane 7), confirming that these transcripts are degraded by the nuclear exosome when nuclear export is defective, consistent with our previous findings [26]. Importantly, overexpression of *RPL22A* in these conditions resulted in a repression of the spliced *RPL22B* transcript (Fig 4B, lane 8). This suggests that the pre-mRNA

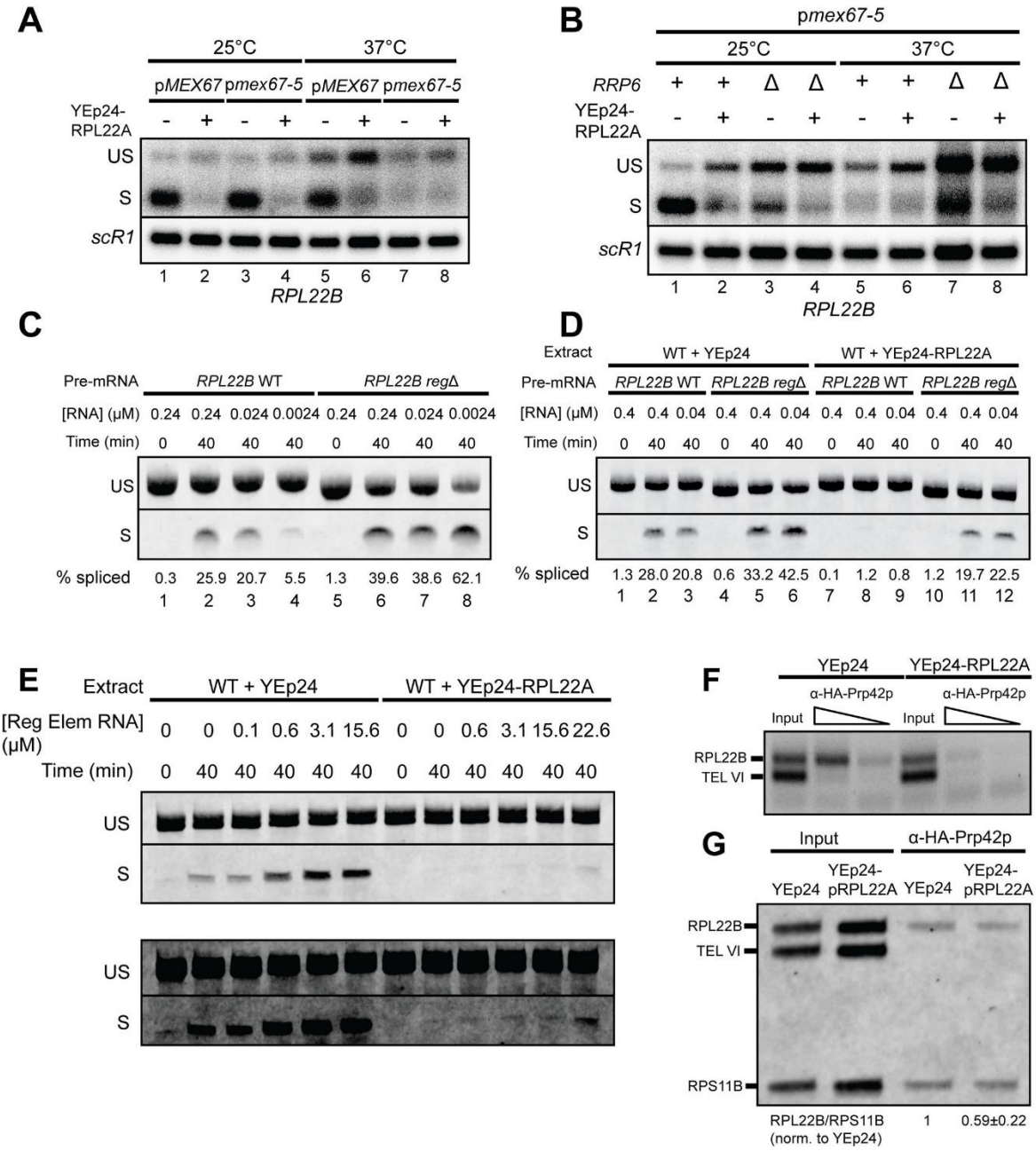


Fig 4. Inhibition of *RPL22B* splicing by Rpl22p in the nucleus can be reproduced *in vitro* and may involve a repression of the co-transcriptional U1 snRNP recruitment *in* vivo. A. Northern blot analysis of *RPL22B* mRNA expression in *mex67*Δ strains carrying plasmids expressing either wild-type *MEX67* or the temperature sensitive *mex67-5* allele and either the empty Yep24 vector or the *RPL22A* overexpression plasmid at permissive (25°C) and non-permissive (37°C) temperatures. The unspliced (US) and spliced (S) RNA species are shown. *SCR1* was used as a loading control. **B.** Northern blot analysis of the *mex67-5* strain similar to panel A with the addition of the genetic inactivation of *RRP6*. Detection methods and transcript labels are similar to panel A. **C.** RT-PCR analysis of *in vitro* splicing reactions in wild-type splicing extract. Decreasing amounts of *RPL22B* wild-type and *reg*Δ pre-mRNA were added to splicing reactions and

incubated at 25°C for 40 minutes. The RNA was then purified and analyzed by RT-PCR. Labeled bands indicate the unspliced (US) and spliced (S) species. The gel image has been truncated for space considerations; the uncut image can be viewed in S13 Fig. **D.** RT-PCR analysis of splicing reactions utilizing wild-type and *RPL22A* overexpression splicing extracts. *RPL22B* wild-type and *regΔ* pre-mRNA were added to splicing reactions at the indicated concentrations and incubated at 25°C for 40 minutes. The RNA was then purified and analyzed by RT-PCR. Labeled bands indicate the unspliced (US) and spliced (S) species. The gel image has been truncated for space considerations; the uncut image can be viewed in S14 Fig. Note also the slightly smaller size of the *regΔ* pre-mRNA versus wild-type due to the deletion of 14 nucleotides. **E.** RT-PCR analysis of *RPL22B in vitro* splicing reactions in the presence of short RNA transcripts containing the regulatory element with light (upper panels) and dark (lower panels) exposures. Splicing extracts were prepared from wild-type yeast carrying empty YEp24 or an *RPL22A* overexpression vector. All splicing reactions contained 0.005 μM *RPL22B* pre-mRNA. Increasing amounts of regulatory element RNA were added to splicing reactions to the final concentrations shown. Splicing reactions were incubated at 25°C for 40 minutes, after which the RNA was purified and used for RT-PCR. Labeled bands indicate the unspliced (US) and spliced (S) species. See S1 Table for primers used to generate the regulatory element RNA transcript. The gel images have been truncated for space considerations; the uncut images can be viewed in S7 Fig. **F.** Analysis of U1 snRNP recruitment at the *RPL22B* locus by chromatin immunoprecipitation. Ethidium bromide stained gel showing PCR products amplified from chromatin obtained from Prp42-HA ChIP or the corresponding input chromatin. The ChIP chromatin concentration was titrated for analysis. DNA was amplified using a multiplexed PCR reaction containing primer sets for the intronic regions of *RPL22B* (target gene) and telomeric region of chromosome VI (TEL VI, ChIP negative control). **G.** Same as E, but Cy3-labeled PCR products were amplified from chromatin obtained from Prp42-HA ChIP or the corresponding input chromatin. DNA was amplified using a multiplexed PCR reaction containing primer sets for the intronic regions of *RPL22B* (target gene) and *RPS11B* (ChIP positive control) and the telomeric region of chromosome VI (ChIP negative control). The quantification of repression of Prp42p association with *RPL22B* chromatin upon overexpression of *RPL22A* in the ChIP samples is displayed below the gel image. Within each ChIP sample the *RPL22B* PCR signal was normalized internally to the *RPS11B* PCR signal. These ratios were then normalized to the ratio for YEp24, providing a measurement of the decrease of Prp42p association with *RPL22B* upon *RPL22A* overexpression. The value shown indicates mean ± 1 standard deviation for two biologically independent ChIP experiments.

doi:10.1371/journal.pgen.1005999.g004

is subject to autoregulation even when it is localized to the nucleus and that the regulatory mechanism functions via inhibition of splicing.

To further confirm that the autoregulatory mechanism functions at the level of splicing regulation, we investigated whether inhibition of *RPL22B* pre-mRNA splicing could be recapitulated *in vitro*, where the transcript is continuously exposed to the splicing machinery. A wild-type version of the *RPL22B* unspliced transcript and a mutant lacking intronic nucleotides 178 through 191 (*regΔ*), which removes a section of the stem necessary for autoregulation and mRNP binding *in vivo* (see above) were generated by *in vitro* transcription. We assembled *in vitro* splicing reactions by adding these pre-mRNA substrates at various concentrations to splicing extracts derived from wild-type cells, and we determined the extent of splicing of the exogenous *RPL22B* by RT-PCR (Fig 4C and 4D). Strikingly, we found that the splicing efficiency of the wild-type pre-mRNA substrate in wild-type extracts was dependent on the concentration of the pre-mRNA, with the proportion of spliced mRNA decreasing progressively as the concentration of the splicing substrate was reduced (Fig 4C, lanes 1–4). We hypothesized that this might be due to the saturation of the relatively small amount of free Rpl22p present in the extracts by the pre-mRNA at higher substrate concentrations, leaving the majority of the pre-mRNA available to be acted upon by the spliceosome; at lower pre-mRNA substrate concentrations there is sufficient free Rpl22p in the splicing reaction to efficiently inhibit splicing. In agreement with this hypothesis, the splicing of the *regΔ* substrate remained efficient with decreasing substrate concentration (Fig 4C, lanes 5–8), also demonstrating that splicing inhibition does not occur in the absence of the regulatory element *in vitro*. To confirm that excess Rpl22p can inhibit the splicing of *RPL22B in vitro*, we analyzed the splicing of both wild-type and *regΔ* substrates in splicing extracts derived from cells overexpressing Rpl22p. As predicted, the overexpression of Rpl22p resulted in a nearly complete suppression of splicing of the wild-type pre-mRNA substrate *in vitro*. (Fig 4D, lanes 7–9). Importantly, deletion of the regulatory element restored splicing to a level comparable to that of the wild-type substrate in wild-type splicing extracts (Fig 4D, compare lanes 1–3 with lanes 10–12). Thus, the intronic regulatory element enables Rpl22p to effectively inhibit the splicing of the *RPL22B* pre-mRNA *in vitro*, and the overall efficiency of splicing is influenced by the relative levels of pre-mRNA and Rpl22p protein. Moreover, splicing of the pre-mRNA is inhibited *in vitro* when the spliceosome cannot be spatially separated from the substrate by cellular relocalization, arguing against a model involving the expedited export of the pre-mRNA to the cytoplasm *in vivo*.

The addition of a short RNA fragment harboring the regulatory element sequences enhances *RPL22B* pre-mRNA splicing *in vitro*

The transposition of the *RPL22B* intronic regulatory element into the *RPS21A* intron did not confer splicing-based autoregulation of the *RPS21A* pre-mRNA (S6 Fig), possibly because this RNA structure does not fold properly in the context of the foreign intron and therefore does not bind Rpl22p. However, we hypothesized that a shorter RNA transcript containing only the regulatory element sequence might effectively sequester Rpl22p, resulting in increased splicing efficiency *in vitro*. To investigate this further, we repeated the *in vitro* splicing reactions utilizing low concentrations of *RPL22B* pre-mRNA, conditions which were shown above to result in low splicing efficiency (Fig 4C). We then added increasing concentrations of a 147 nucleotides-long RNA transcript harboring the regulatory element to test whether the addition of this transcript would enhance pre-mRNA splicing. This experiment showed that the regulatory element RNA sequence alone is sufficient to increase splicing of the pre-mRNA in a concentration-dependent manner (Figs 4E and S7). Importantly, splicing reactions in extracts derived from yeast overexpressing Rpl22p required comparably higher concentrations of the regulatory element RNA in order for splicing efficiency to be improved (Figs 4E and S7), further showing that splicing efficiency is directly influenced by the amount of Rpl22p protein unbound to the regulatory element. Taken together, these results strongly suggest that, at least *in vitro*, the isolated regulatory element is sufficient to directly or indirectly bind and sequester Rpl22p so that the protein is unable to inhibit splicing of the *RPL22B* pre-mRNA.

Rpl22p overexpression reduces co-transcriptional recruitment of U1 snRNP to the *RPL22B* locus

Previous chromatin immunoprecipitation (ChIP) studies have shown that *RPL22B* is one of many intron-containing chromosomal loci to which the U1 snRNP is recruited co-transcriptionally [39]. In light of these findings, we hypothesized that the repression of *RPL22B* pre-mRNA splicing may be due in part to the interference of co-transcriptional U1 recruitment. To test this, we performed ChIP on yeast strains expressing HA-tagged versions of the U1 snRNP protein Prp42p to determine whether the overexpression of Rpl22p impedes the association of the protein with chromatin at the *RPL22B* locus. Indeed, we found that yeast cells carrying the *RPL22A* overexpression vector exhibited a slightly repressed enrichment of Prp42p at the *RPL22B* locus relative to cells transformed with the control vector (Fig 4F). As a normalization control, we also examined the recruitment of Prp42p to *RPS11B*, another RPG which demonstrates strong co-transcriptional U1 recruitment [39] and should be relatively unaffected by Rpl22p overexpression (Fig 4G). Quantitative analysis showed that Rpl22p overexpression reduced U1 snRNP co-transcriptional recruitment by about 40%. These experiments suggest that Rpl22p inhibits U1 snRNP recruitment to the *RPL22B* locus. We carefully note, however, that this result only gives a partial explanation for the mechanism of splicing inhibition, as the repressive effect of Rpl22p overexpression on *RPL22B* pre-mRNA splicing *in vivo* and *in vitro* (Figs 1D–1G, 4D and 4E) appears more substantial than the minor repression of U1 recruitment detected here.

Rpl22 paralog identity does not dictate cellular sensitivity to ribosome inhibitors

In wild-type yeast cells the proportion of ribosomes harboring the Rpl22A paralog exceeds the proportion harboring the Rpl22B paralog by a factor of more than 15 [18], suggesting that the *A* paralog plays the frontline role for actively-translating ribosomes. The dearth of Rpl22Bp in wild-type ribosomes begs the question of whether yeast have evolved to specifically regulate the

RPL22 paralog composition of ribosomes, perhaps to influence cell fitness when challenged with antagonistic conditions such as chemical stressors. We therefore performed an antibiotic screen utilizing an array of chemicals known to target the ribosome, including anisomycin, paromomycin, puromycin, neomycin, and verrucarin A [40,41]. We assessed the growth of wild-type and *rpl22aΔ* strains transformed with an empty vector or with plasmids overexpressing either intronless *RPL22A* or *RPL22B*. In general, growth of wild-type cells was unimpeded by the antibiotic concentrations tested and was not affected by the overexpression of either paralog (S10 Fig). The *rpl22aΔ* strain demonstrated the slowest growth in nearly every condition. However this growth defect could be rescued by the overexpression of either *RPL22* paralog, except in the case of verrucarin A (S10 Fig). Interestingly, *rpl22aΔ* demonstrated more robust growth in the presence of this chemical compared to the wildtype, possibly due to a defect in ribosomal subunit assembly that may confer hyperresistance to this antibiotic [40]. Overall, our observations do not support the notion of Rpl22p paralog composition influencing cellular sensitivity to the particular ribosome inhibitors and concentrations tested, though we note that our screen was limited in scope and certainly does not preclude this possibility.

Inhibition of *RPL22B* pre-mRNA splicing promotes down-regulation of mature mRNA during MMS-Induced DNA damage

Splicing-specific microarray studies have revealed that ribosomal protein genes undergo changes in pre-mRNA processing in response to a variety of environmental stresses, including amino acid starvation, growth in non-fermentable carbon sources, osmotic stress and heat shock [27,42]. Earlier experiments have also shown that the transcript levels of many RPGs are repressed by exposure to methane methyl-sulfonate (MMS) [1], an alkylating agent that is thought to stall DNA replication fork movement. In agreement with these results, we observed a tapering of spliced *RPL22B* mRNA levels in cells treated with 0.05% MMS (Fig 5A). Surprisingly, MMS-treatment of the NMD-deficient strain *upf1Δ* resulted in a substantial accumulation of *RPL22B* pre-mRNA that persisted even as the mature transcript gradually decreased (Fig 5A). Based on this result, we hypothesized that the repression of mature *RPL22B* transcript during MMS-induced DNA damage relies primarily on the inhibition of splicing rather than transcriptional repression, which may in turn depend on the regulatory mechanism that we identified herein.

To test whether the *RPL22B* intronic regulatory element regulates the splicing of *RPL22B* pre-mRNA during DNA damage, we treated a strain harboring a chromosomal deletion of the intronic regulatory element (*RPL22Bi regΔ*) with 0.05% MMS for three hours and compared its splicing patterns to those of the wild-type strain. As described above, the wild-type strain exhibited a persistent accumulation of unspliced transcript as the mature transcript decreased over the course of the treatment. By the end of the treatment the proportion of unspliced transcript was noticeably higher than the proportion of spliced transcript (Fig 5B). However, removal of the regulatory element largely negated this effect—the pre-mRNA did not accumulate over the three-hour treatment and the spliced transcript remained the predominant species throughout the time course of the experiment (Fig 5B). This difference was also apparent when comparing quantitatively the reduction of the amount of spliced transcript triggered by the MMS treatment within each strain. Whereas a 3 hour treatment reduced spliced *RPL22B* to c.a. 18% of its pre-treatment levels in the wild-type strain, deletion of the regulatory element resulted in an attenuated reduction to only 58% of its pre-treatment levels (S11A Fig). Thus, the inhibition of splicing through the *RPL22B* intronic regulatory element appears to play a major role in maximizing the cell's ability to suppress levels of mature *RPL22B* transcript during DNA damage.

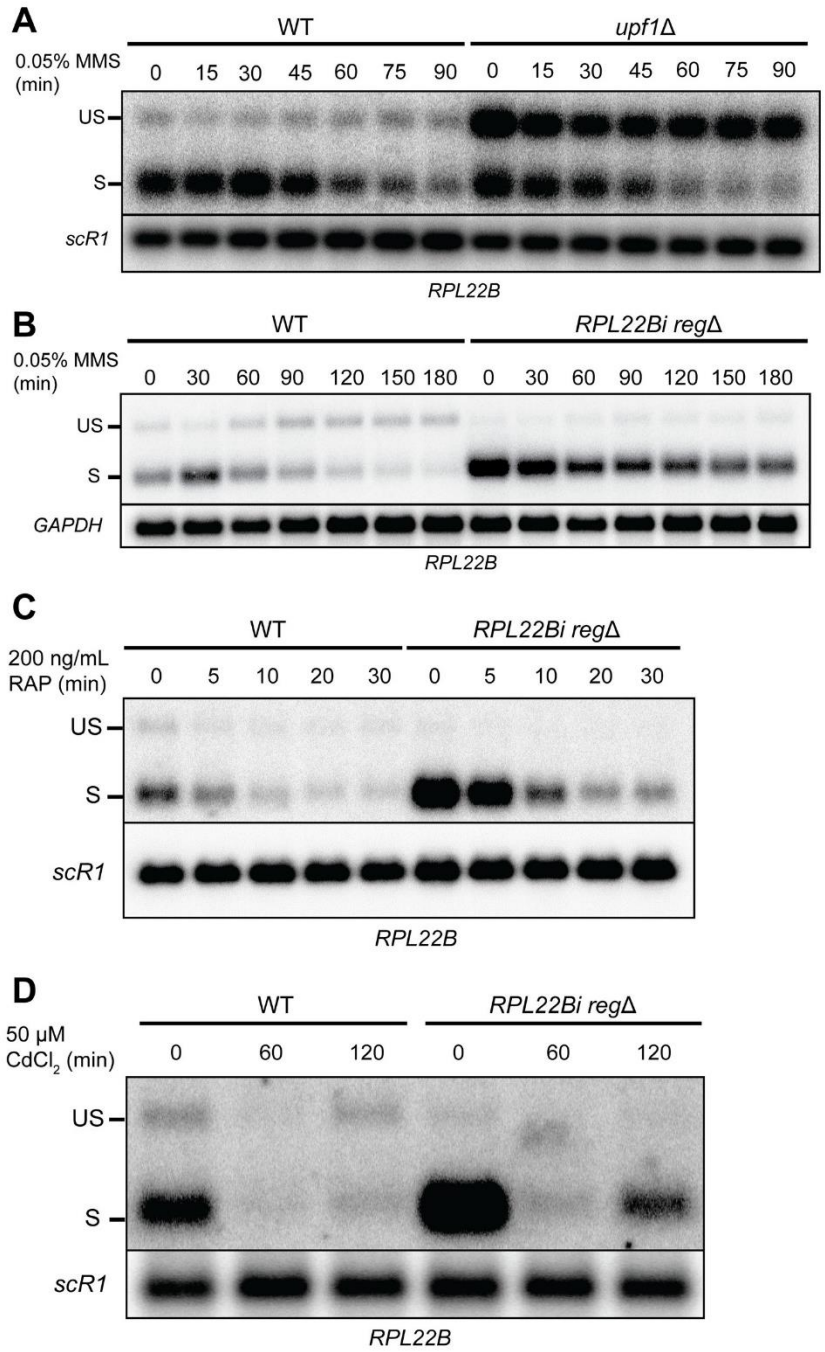


Fig 5. Inhibition of Rpl22p-mediated splicing autoregulation attenuates downregulation of *RPL22B* in response to MMS-induced DNA damage but not rapamycin treatment. A. Northern blot analysis of splicing kinetics for *RPL22B* in wild-type and *upf1*Δ strains in response to MMS-induced DNA damage. Cells were grown to exponential phase in YPD media and then shifted to YPD media containing 0.05% MMS for 90

minutes. *RPL22B* transcripts were detected with an *RPL22B* 5'UTR riboprobe. Shown are the unspliced (US) and spliced (S) species. *SCR1* was used as a loading control. **B.** Northern blot analysis of splicing kinetics for *RPL22B* in wild-type yeast and in a strain in which the regulatory element has been removed from the *RPL22B* intron (*RPL22Bi regΔ*). Cells were grown in YPD containing 0.05% MMS for 180 minutes. Bands are labeled similarly to panel A. *GAPDH* was used as a loading control. **C.** Northern blot analysis of splicing kinetics for *RPL22B* in wild-type yeast and in a strain in which the regulatory element has been removed from the *RPL22B* intron (*RPL22Bi regΔ*). Cells were grown in YPD containing 200 ng/mL rapamycin for 30 minutes. Detection methods and labels are similar to panels A-B. **D.** Northern blot analysis of *RPL22B* splicing in WT and *RPL22Bi regΔ* strains when were grown in YPD containing 50 μM $CdCl_2$ for 120 minutes. Detection methods and labels are similar to panels A-C.

doi:10.1371/journal.pgen.1005999.g005

To determine whether the Rpl22 protein is necessary for regulation in DNA damage conditions, we repeated the MMS treatment experiment in the *rpl22aΔ* strain. In the wild-type strain, we observed a shift in splice patterns in which the levels of unspliced transcript becomes comparable to the spliced species after three hours of MMS treatment (S11B Fig). By contrast, removal of the intronic regulatory element (*cis*, *RPL22Bi regΔ*) or deletion of *RPL22A* (*trans*, *rpl22aΔ*) prevented this shift and resulted in the spliced species remaining the predominant form (S11B Fig). These results demonstrate that the down-regulation of mature *RPL22B* transcript levels depends at least in part on the ability to inhibit splicing of the pre-mRNA through the intronic regulatory element.

To test whether the down-regulation of either of the mature *RPL22* transcripts impacts cell growth in these conditions, wild-type yeast carrying an empty vector or overexpression plasmids for either *RPL22A* or *RPL22B* were monitored for cell growth in a range of MMS concentrations. Importantly, because the exogenous *RPL22* transcripts are intronless, they are unaffected by the inhibitory mechanism. Cell growth was similarly impeded by MMS treatment in a concentration-dependent manner for all three strains (S11C Fig), suggesting that the inability to repress of this specific protein *per se* is not detrimental to cell fitness when grown in the presence of MMS. However, we propose that the biological significance of *RPL22B* pre-mRNA splicing inhibition in MMS-induced DNA damage may be rooted in the liberation of the spliceosomal machinery rather than the repression of the *RPL22B* transcript itself (see Discussion).

Rapid changes in yeast pre-mRNA levels are known to occur not only in MMS-induced DNA damage but also in a variety of environmental stress conditions. For example, amino acid depletion results in the inhibition of splicing and a subsequent accumulation of RPG pre-mRNAs [42]. Additionally, inhibition of the target of rapamycin complex 1 (TORC1) via rapamycin treatment has been shown to suppress RPG expression by preventing the phosphorylation of effector proteins that modulate RPG transcription [43]. *RPL22B* was found to be one of many RPGs whose transcript levels are decreased upon rapamycin treatment [27]. For these reasons we investigated whether the suppression of this transcript during TORC1 inhibition uses a mechanism akin to the autoregulation of splicing described during MMS-induced DNA damage. To address this question, we analyzed the relative levels of spliced and unspliced *RPL22B* transcript in WT and *RPL22Bi regΔ* cells after rapamycin treatment. As expected, the levels of spliced transcript were reduced in both strains following rapamycin treatment (Fig 5C). However, in contrast to what was observed during MMS treatment, the unspliced transcript did not accumulate during rapamycin treatment and the spliced transcript remained the dominant species (Fig 5C). Thus, the regulation of *RPL22B* in the presence of rapamycin appears to rely primarily on transcriptional repression, whereas down-regulation during MMS-induced DNA damage primarily involves auto-inhibition of pre-mRNA splicing.

Disruption of rRNA transcription causes inhibition of *RPL22B* pre-mRNA splicing

The observation that *RPL22B* pre-mRNA splicing appeared to be inhibited upon genetic inactivation of *RPL31A* and *RPL39* (Fig 1A) suggests that this effect might be due to defective ribosome assembly in these mutants. A reduction in the efficiency of ribosome assembly would increase the population of free ribosomal proteins or pre-ribosomal protein complexes, resulting in excess amounts of free or complexed Rpl22p available to inhibit the splicing of the pre-mRNA. Similar results might be expected in other conditions that impede the assembly of ribosomes, including the down-regulation of rRNA transcription. To test this hypothesis, we took advantage of the ability of cadmium to suppress rRNA transcription by RNA polymerase I [44]. Treatment of wildtype and *RPL22Bi reg*Δ strains with 50 μM $CdCl_2$ for two hours resulted in a reduction in the levels of 35S rRNA as well as other intermediates in the rRNA biogenesis pathway (S12 Fig). Following an initial decrease of all *RPL22B* transcripts at 60 minutes, analysis of *RPL22B* transcripts at 120 minutes revealed a clear inhibition of pre-mRNA splicing (Fig 5D, lanes 1 through 3). This inhibition was not observed in the absence of the regulatory element, resulting in the spliced species remaining the predominant form after 120 minutes (Fig 5D, lanes 4 through 6), similar to the behavior observed during MMS treatment. Thus, the disruption of rRNA transcription can inhibit *RPL22B* pre-mRNA splicing, presumably by suppressing *de novo* ribosome biogenesis and increasing cellular concentrations of free or complexed Rpl22p. Overall, these results show that autoregulation of *RPL22B* splicing is a major mechanism that contributes to the downregulation of *RPL22B* in conditions which are known to decrease ribosome biogenesis.

Discussion

In this study, we have shown that *S.cerevisiae* utilizes an autoregulatory system to control the cellular levels of the *RPL22B* mRNA. The Rpl22 protein directly or indirectly recognizes and binds to an RNA structure present in the *RPL22B* intron, resulting in inhibited splicing of the pre-mRNA. Deletion of the regulatory element results in a substantial increase in the levels of spliced *RPL22B* mRNA at steady-state, indicating that *RPL22B* is subject to continuous repression by its own protein product. Altogether, our results illustrate a model in which the binding of Rpl22p to the *RPL22B* pre-mRNA interferes with splicing of the transcript. The pre-mRNA is then susceptible to degradation by exonucleases in the nucleus or to export and 5'-to-3' degradation by Xrn1p, likely as the final step of the NMD pathway (Fig 6). This regulatory activity constitutes an extra-ribosomal function for Rpl22p that serves to regulate the cellular levels of the protein and, by extension, the Rpl22p paralog composition of ribosomes.

Because we analyzed Rpl22p binding to the *RPL22B* pre-mRNA *in vivo*, our experiments did not determine whether Rpl22p binds the pre-mRNA as a free protein or as part of a larger RNP complex, including possibly the entire pre-60S LSU, the latter being possible due to Rpl22p's external position on the ribosome [45]. However, because *RPL22* overexpression led to a decrease in pre-mRNA splicing presumably without concomitant stoichiometric increases in the expression of other RPGs, we speculate that the binding to the pre-mRNA could involve either free protein or in association with other proteins, but likely not the entire LSU. Our results also revealed inhibition of pre-mRNA splicing following the disruption of rRNA transcription (Fig 5D), presumably by impeding ribosome assembly and relinquishing free or complexed Rpl22p. In addition to Rpl22p, this model would also predict that the extra-ribosomal functions of other ribosomal proteins might become more prevalent in these conditions.

Finally, our results suggest that the co-transcriptional recruitment of the U1 snRNP to the *RPL22B* locus is at least partially inhibited by the increased presence of Rpl22p, which may

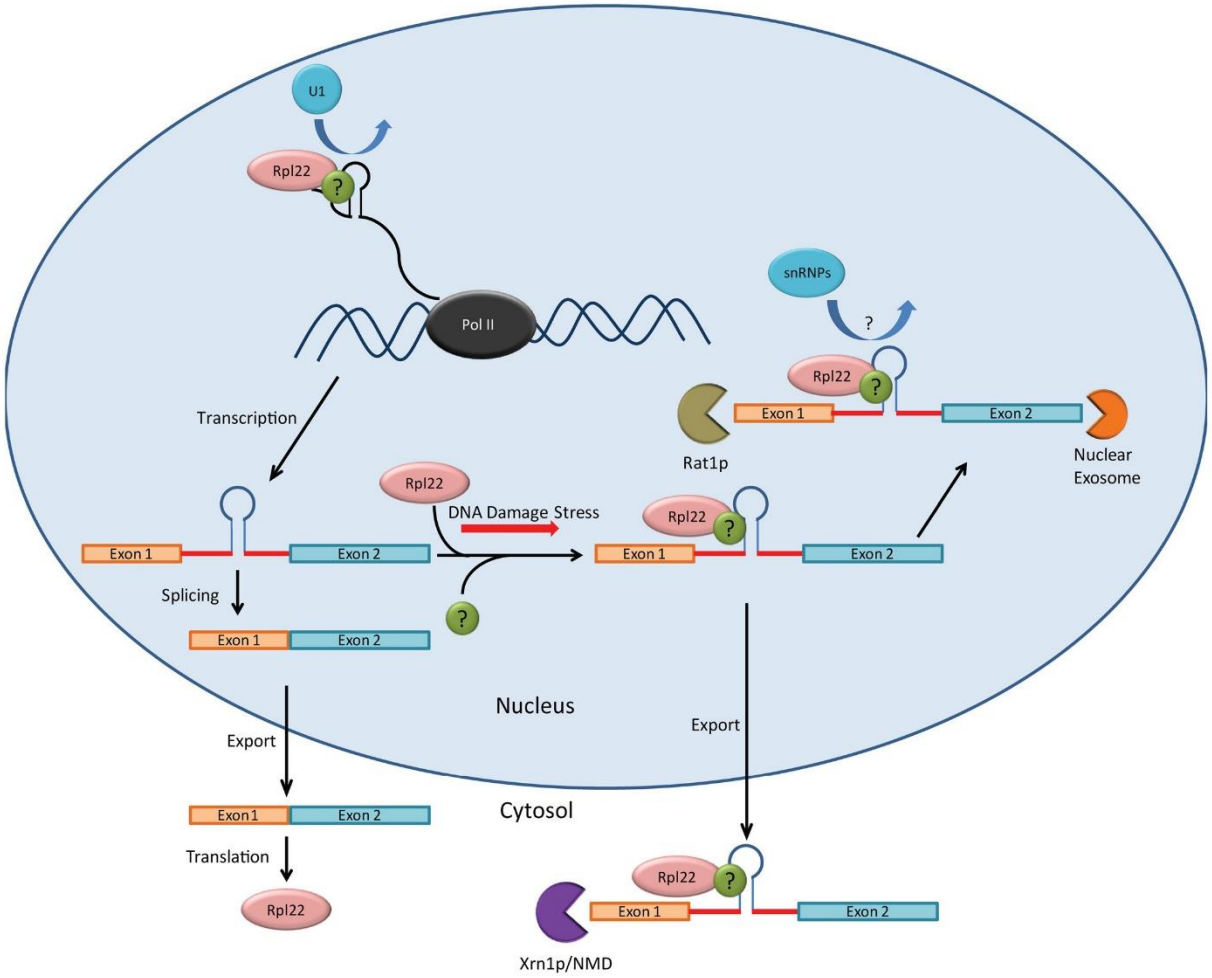


Fig 6. Model for the autoregulation of *RPL22B*. *RPL22B* pre-mRNA contains an intronic regulatory element, depicted here by a simplified stem loop, to which the Rpl22 protein is able to directly or indirectly bind to associate with the unspliced mRNP. This association inhibits splicing of the pre-mRNA, potentially by prohibiting binding of one or more spliceosomal snRNPs, which is then either degraded by nuclear exonucleases or exported to the cytoplasm and degraded by NMD. Splicing inhibition is exacerbated during DNA damage stress. The green protein labeled with the black question mark represents potential binding factors that may facilitate an indirect interaction between Rpl22p and the *RPL22B* unspliced mRNP.

doi:10.1371/journal.pgen.1005999.g006

explain some of the loss of splicing of the *RPL22B* pre-mRNA (Fig 4F and 4G). However, we are careful to note that this mechanism may not fully account for the entirety of the splicing inhibition. In addition, *in vitro* splicing experiments suggest that splicing can also be inhibited post-transcriptionally (Fig 4C and 4D), in agreement with evidence suggesting that the splicing of *S. cerevisiae* pre-mRNAs containing short second exons (<500 nt; the second exon of *RPL22B* is 377 nt in length) is primarily post-transcriptional [46]. We speculate that post-transcriptional splicing inhibition may entail the prevention of one or more of the spliceosomal snRNPs from effectively binding the pre-mRNA (Fig 6).

The intronic RPG regulation study conducted by Parenteau and colleagues found that the choromosomal deletion of the *RPL22A* intron resulted in a >10-fold repression of *RPL22B* mRNA levels and a 75% increase in *RPL22A* mRNA levels [4]. In agreement with these findings, our overexpression experiments with intronless *RPL22A* likewise resulted in a substantial repression of mature *RPL22B* (Fig 1). Furthermore, the chromosomal deletion of the *RPL22B* intron triggered a >6-fold increase in mature *RPL22B* transcript levels [4]. Based on our results, we can confidently attribute this observation to the role of the *RPL22B* intron in regulating the expression of the mature transcript through the splicing-mediated inhibitory mechanism that we have identified herein. Thus, our results are very consistent with previous observations and provide additional mechanistic insight into the regulatory activity of *RPL22*.

A unique regulatory element for a familiar purpose

The structure and composition of the *RPL22B* regulatory element and the associated regulatory mechanism that we have identified in this study is reminiscent of other cases of ribosomal protein autoregulation, some of which were shown to be based on splicing. A classical example of a yeast RPG autoregulatory system was identified for *RPL30* by Eng and Warner [8]. In this case, the regulatory element pairs the nucleotides of the intronic 5' splice site with nucleotides further upstream on the pre-mRNA transcript, thereby sequestering the 5' splice site into a secondary structure. The Rpl30 protein stabilizes this configuration by binding to an internal loop structure immediately adjacent to the 5'-splice site, thereby preventing the U2 snRNP from accessing the substrate [47,48]. A similar mechanism has been elucidated for *RPS14B*, in which the Rps14 protein was found to bind the *RPS14B* pre-mRNA in which the 5' splice site was likewise predicted to fold into a secondary structure, repressing its expression [9], although the question of whether this directly impacts pre-mRNA splicing was not addressed. Other studies focusing on the mechanisms of 3'SS selection have found that that the availability of this splice site, and therefore splicing propensity, is largely dictated by its sequestration into pre-mRNA secondary structures, which in turn is influenced by temperature [49]. By contrast, the regulatory element identified herein for *RPL22B* does not include any of the three primary sites required for spliceosome assembly, arguing against a direct "shielding" of splice sites from access by the spliceosome. However, we cannot rule out the possibility that the formation of this element impacts the folding of the remainder of the intron, which may in turn dictate the inclusion of these sites within more complex structures. Alternatively, it is possible that the binding of the protein induces a conformational change in the *RPL22B* mRNP that disfavors splicing by steric hindrance. Thus, the precise mechanism by which direct or indirect binding of Rpl22p to the intronic regulatory element inhibits splicing remains to be investigated. We could not demonstrate that the regulatory element was sufficient to confer regulation *in vivo* when transposed into a different intron. However, the regulatory element was sufficient to sequester Rpl22p and therefore enhance *RPL22B* pre-mRNA splicing *in vitro* when added *in trans*. These results suggest that the regulatory element may not fold properly within the context of a foreign intron, but is functional when introduced in isolation, at least in *vitro*.

An Rpl22 autoregulatory system has previously been identified in higher eukaryotes. Based on previous knowledge that human Rpl22 binds strongly to Epstein-Barr virus-expressed RNA 1 (EBER1), an *in vitro* search for Rpl22 RNA ligands revealed a consensus motif characterized by an RNA stem closed with a G-C base pair that leads into a hairpin loop consisting of between 5 to 9 nucleotides, of which the 3'-most nucleotide is a uridine [36]. Accordingly, a recent study by O'Leary and colleagues identified an RNA-binding motif matching these characteristics within exon 2 of the zebrafish *Rpl22l1* mRNA that allows Rpl22 protein to bind and regulate the expression of the mRNA. They further confirmed that this binding is lost upon

disruption of the hairpin loop [13]. In the yeast system, however, the regulatory element is entirely intronic and the fidelity of binding appears to be conferred by a combination of sequence-independent secondary structures and sequence-specific elements within the stem region. Specifically, our mutational analyses showed that the identity of the nucleotides within the apical hairpin loop of the *RPL22B* regulatory element is inconsequential in determining binding specificity (S1 Text and S4 Fig). It is also uncertain whether there is mechanistic overlap with the mammalian and zebrafish autoregulatory systems, specifically whether these mechanisms operate through inhibition of splicing as we have demonstrated to be the case in yeast. Ostensibly, however, the yeast and metazoan Rpl22 regulatory mechanisms appear to have the identical purpose of repressing one of the two Rpl22 paralogs. It remains to be seen whether the propensity to down-regulate Rpl22 paralogues has resulted from evolutionary conservation among eukaryotes, or if autoregulatory mechanisms have emerged as a result of convergent or parallel evolution.

Why regulate *RPL22B*?

Although autoregulatory mechanisms exist for multiple ribosomal protein genes in yeast, it is not immediately clear what purpose they serve, if any, with regards to cell's ability to respond to environmental stresses. A previous study found that ribosomal reprogramming and enrichment with the Rpl22Ap paralog occurs in yeast following exposure to oxidative stress, though it remains uncertain how this reprogramming ultimately affects the proteome [18]. Here we uncovered a role for the Rpl22 autoregulatory system in preventing the splicing of *RPL22B* pre-mRNA in conditions of MMS-induced DNA damage (Fig 5). At this juncture, we have not determined whether this inhibition is a direct effect of MMS treatment or is instead a downstream result of the DNA damage response. However, given that ribosome assembly is typically disrupted during DNA damage [50,51], it is possible that the MMS treatment triggers increased levels of free or complexed Rpl22p that are able to inhibit *RPL22B* splicing, perhaps similar to cadmium treatment (Figs 5D and S12). Because the overexpression of *RPL22B* cDNA did not result in any noticeable differences in cellular fitness compared to the wild-type cells when grown in a range of MMS concentrations (S11 Fig), we propose that, at least in yeast, the regulation of this specific protein is not essential to maximize cell survival in these conditions. Rather, it is more likely part of a broader gene repression program designed to respond to DNA damage, and is perhaps a result of the documented role of the spliceosome in promoting decay of the *BDF2* mRNA during DNA damage response [52]. Additionally, the availability of spliceosomal components is a key limiting factor in determining the efficiency of splicing; pre-mRNA transcripts "compete" for spliceosome components in situations when there is limited spliceosome available [53]. In this regard, the Rpl22 autoregulatory mechanism may help prevent the sequestration of the spliceosomal machinery on a non-essential transcript, leaving more spliceosomes available for the processing of more crucial transcripts, which may prove vital in cases of cellular stress. The additional observation that *RPL22B* down-regulation during rapamycin treatment does not appear to involve inhibition of splicing (Fig 5C) may further suggest that the mechanism features more prominently under specific stress conditions than in others, the reasons for which will be interesting inquiries for future experiments.

In mammalian cells Rpl22 has been shown to play a critical role in the maturation of αβ lineage T cell development as well as in early B cell development, both of which are dependent upon the induction of the p53 tumor suppressor [54,55]. Ablation of Rpl22 results in activation of p53 and the developmental arrest of both cell types in a tissue-specific manner. Interestingly, p53 is also activated in response to DNA damage, one of the early hallmarks of tumorigenesis [56]. While it remains to be determined whether Rpl22 expression is affected by DNA damage in

mammalian cells as it is in yeast, it is tempting to speculate that the ability for yeast cells to down-regulate *RPL22B* during MMS-induced DNA damage may be a functional progenitor that evolved to modulate the induction of p53 in response to DNA damage in higher eukaryotes.

In summary, our study reveals an additional splicing-based autoregulatory pathway for a specific yeast RPG. In light of recent genome-wide analyses revealing numerous cases of negative feedback of RNA expression that serves to buffer allelic variants in yeast [57], we propose that additional mechanisms are in place to assist the cell with genetic autoregulation at many of these loci. Additional studies will delve into these mechanisms and help to demystify the intricacies of eukaryotic gene expression control.

Materials and Methods

Strain construction

Unless otherwise indicated, all mutant strains in this study were derived from the BY4742 background (S1 Table) or obtained directly from the GE Dharmacon Yeast Knockout Collection. Deletion mutants created in this study were constructed by transformation of the background strain with PCR products containing an auxotrophic marker or antibiotic resistance cassette as well as sufficient flanking sequences for efficient homologous recombination into the ORF of interest (see below). Successful transformants were screened by growth on dropout media or antibiotic plates and confirmed by PCR. Non-deletion chromosomal mutations including the insertion of the Myc tag into the chromosomal *RPL22A* locus, replacement of the *RPL18B* intron with the *RPL22B* intron, and deletion of the regulatory element from the *RPL22B* intron were achieved using the *delitto perfetto* approach [32]. Briefly, the CORE cassette was amplified from the pCORE plasmid by high-fidelity PCR (New England Biolabs No. M0530) using primer sets containing homologous sequences for the regions named above. CORE cassettes were transformed into yeast strains using the LiAc/SS carrier DNA/PEG approach [58]. Transformants were plated onto YPD and then replica plated one day later onto G418 plates. G418-resistant colonies indicating successful CORE integration were confirmed by PCR using flanking and internal primers. Next, for the Myc insertion and intron replacement mutations, high-fidelity PCR products were generated for the 13x Myc tag or the *RPL22B* intron containing flanking regions at the *RPL22A* or *RPL18B* ORFs, respectively. A total volume of 400 μl of PCR products was added to Eppendorf tubes containing 1 ml 100% EtOH and 40 μl 3 sodium acetate pH 5.2, precipitated for ≥30 minutes at -80°C, then centrifuged for 15 minutes at 15000 rpm to pellet the DNA. Pellets were washed with 500 μl 70% EtOH, centrifuged for 5 minutes, air dried and resuspended in 34 μl nuclease-free water. For the chromosomal deletion of the regulatory element from the *RPL22B* intron, overlapping oligonucleotides were used containing the deletion of interest and flanking regions for the *RPL22B* intron. See S1 Table for a full list of oligonucleotides used for these procedures. For all mutations, DNA was transformed into CORE-containing strains as described above. Transformants were plated on YPD and then replica plated onto 5-FOA plates one day later to screen for successful CORE excision. Colonies growing on 5-FOA were tested for G418 sensitivity to confirm CORE excision. Finally, genomic DNA was extracted from screened clones (see below) and mutations were confirmed by PCR and Sanger sequencing (Laragen, Inc). The *rpl31aΔ* and *rpl39Δ* strains were generously provided by Brian Kennedy (Buck Institute). The HA-tagged Prp42p strains were generously provided by Tracy Johnson (UCLA).

Yeast cell culturing

For experiments involving untransformed cells, cultures were grown in rich media (YPD: 1% w/v yeast extract, 2% w/v peptone, and 2% w/v dextrose). Cells transformed with pUG23 or its

derivative plasmids were grown in–HIS-MET synthetic dropout media (0.67% w/v yeast nitrogen base, 2% w/v dextrose, and 0.2% w/v powdered amino acid mixture without histidine and methionine). Cells transformed with pUG35, YEp24 or their derivatives were grown in–URA (for YEp24) or–URA-MET (for pUG35) synthetic dropout media (0.67% w/v yeast nitrogen base, 2% w/v dextrose, and 0.2% w/v powdered amino acid mixture without uracil or without uracil and methionine). For steady-state analysis experiments, cell cultures were grown to exponential phase until an OD_{600} of ~0.4–0.6 was attained. Cultures were then harvested in 50 ml volumes and centrifuged for 3 minutes at 3000 rpm (Sigma Rotor 11030), washed with deionized water and transferred to 2 ml screw-cap Eppendorf tubes. Cells were then centrifuged for 30 seconds at 13200 rpm in a benchtop microcentrifuge. Following removal of the supernatant, cells were flash frozen in liquid nitrogen. For MMS treatment experiments, 100% methyl methanesulfonate (Acros Organics No. AC15689-0050) was added to exponentially-growing cells to a final concentration of 0.05%. Cultures were then grown for an additional 90–180 minutes with culture harvests at regular intervals as indicated in the text. For heat shift experiments, exponentially-growing cells were centrifuged as described above and then sterile transferred to pre-warmed media at 37°C and grown for an additional three hours. Cultures were then harvested as described above. For rapamycin treatment experiments, 1 mg/ml rapamycin in 90% ethanol, 10% Tween 20 was added to a final concentration of 200 ng/ml. Cultures were then grown for an additional 30 minutes with harvests at the time points indicated in the text. For cadmium treatment experiments, cadmium chloride was added to exponentially-growing cells to a final concentration of 50 μM. Cultures were then grown for an additional 120 minutes with cultures harvests at 60 and 120 minutes as indicated in the text. See S1 Table for a full list of strains and plasmids used in this study.

Yeast genomic DNA extraction

Freshly-streaked yeast cells were collected from media plates and added a 2 ml screw-cap Eppendorf tubes with 400 μl of genomic DNA lysis buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0), 400 μl of acid-washed glass beads (Sigma No. G8772), and 400 μl of DNA Phenol-Chloroform (25 phenol:24 chloroform: 1 isomyl alcohol, pH 8.0, OmniPur). Samples were vortexed for 3 minutes at maximum intensity and then centrifuged for 5 minutes at 13200 rpm in a benchtop microcentrifuge. The top aqueous phase was transferred to a 1.5 ml microcentrifuge tube containing 400 μl of DNA phenol-chloroform. Samples were vortexed for 1 minute at room temperature and centrifuged at 15000 rpm for 5 minutes. 400 μl of the top aqueous phase were transferred to a 1.5 ml microcentrifuge tube containing 280 μl of 100% isopropanol. DNA was precipitated at -80°C for 30 min. The tubes were then centrifuged for 15 min at 15 krpm to pellet the DNA. After removing the supernatant, DNA pellets were washed with 500 μl of 70% ethanol, centrifuged for 5 min at 15 krpm, and resuspended in 300 μl of nuclease-free water. 1.5 μl of isolated gDNA was used for each 50 μL PCR reaction.

RNA extraction

Cells were placed in 2 ml screw-cap Eppendorf tubes and frozen at -80°C. 350 μl of RNA buffer (50mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM EDTA pH 8.0), 350 μl of RNA buffer + 10% SDS, 700 μl of RNA Phenol-Chloroform (25 phenol:24 chloroform: 1 isomyl alcohol, pH 6.7, OmniPur), and 400 μl of acid-washed glass beads were added to cells. Afterwards, cells were vortexed at 3°C for 1 min, with 2 min intervals on ice for a total of 3 cycles. The tubes were then incubated for 6 minutes at 65°C. After incubation, tubes were vortexed for 1 min at high speed and centrifuged for 5 min at 13.2 krpm. 600 μl of the top aqueous layer was transferred

to a 1.5 ml microcentrifuge tube containing 600 μl of RNA phenol-choloroform. Samples were vortexed for 1 minute at room temperature and centrifuged at 15krpm for 2 minutes. 500 μl of the top aqueous phase was transferred to a 1.5 ml microcentrifuge tube containing 500 μl of RNA phenol-chloroform. Samples were vortexed for 1 minute at room temperature and centrifuged at 15krpm for 2 minutes. Finally, 400 μl of the top aqueous phase were transferred to a 1.5 ml microcentrifuge tube containing 1 ml of 100% ethanol and 40 μl of 3M sodium acetate pH 5.2. Tubes were then cooled at -80°C for 30 min. The tubes were then centrifuged for 15 min at 15 krpm to pellet the RNA. After removing the supernatant, RNA pellets were washed with 500 μl of 70% ethanol, centrifuged for 5 min at 15 krpm, and resuspended in 30–40 μl nuclease-free water.

Preparation of splicing extracts

To prepare splicing extracts, 3 liters of cell culture were grown to $OD_{600} \approx 1.6$ as described above and harvested in 500 ml volumes by centrifugation at 5000 RPM for 10 minutes in a Beckman Coulter Avanti J-20 XP centrifuge with a JLA-8.1 rotor chilled to 4°C. The supernatant was removed and cell pellets were resuspended in 400 ml of ice cold buffer AGK (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 200 mM KCl, 10% glycerol, 0.5 mM DTT) and centrifuged again for 10 minutes at 5000 RPM. The supernatant was removed and the cells were resuspended in 30 ml of cold AGK and transferred to a 50 ml centrifuge tube. The cells were centrifuged for 10 minutes at 3500 RPM in a Beckman Coulter Avanti J-25 centrifuge with a JA-25.5 rotor chilled to 4°C. The supernatant was removed and the cells were resuspended in a 0.4 volume of buffer AGK supplemented with 1x protease inhibitor cocktail (Roche No. 04693132001). Cells were then flash frozen via drip suspension into liquid nitrogen and stored at -80°C. For lysis, frozen cells were added to a 50 ml grinding jar (Retsch No. 014620216) containing a 25 mm grinding ball (Retsch No. 053680105) that were pre-chilled in liquid nitrogen. Cells were milled in a Retsch MM400 mixer mill for 5 alternating 3 minute periods of 11 and 12 Hz. Powdered lysate was transferred to an oakridge centrifuge tube and stored at -80°C. To prepare the extract, the powdered lysate was thawed on ice and centrifuged for 30 minutes at 17000 RPM at 4°C. (Beckman Coulter Avanti J-25 centrifuge with JA-25.5 rotor). The supernatant was then transferred to an ultracentrifuge tube and centrifuged at 37000 RPM for 1 hour at 4°C (Beckman Coulter Optima LE-80K Ultracentrifuge with Type 70.1 Ti rotor). Following ultracentrifugation, the middle aqueous phase was extracted and dialyzed in 3.5K MWCO Slide-A-Lyzer dialysis cassettes (Life Technologies No. 66332) in dialysis buffer (20 mM HEPES, 0.2 mM EDTA, 50 mM KCl, 20% glycerol, 0.5 mM DTT, 0.1 mM PMSF) under gentle agitation for two hours, after which the buffer was replaced and dialysis was performed for another 2 hours. Dialyzed splicing extract was transferred from the cassette to microcentrifuge tubes and centrifuged at 10 krpm for 1 minute at 4°C. The supernatant was then aliquoted into to fresh microcentrifuge tubes, frozen in liquid nitrogen, and stored at -80°C until needed.

In vitro splicing assays

Pre-mRNA templates were transcribed using the MEGAscript T3 Transcription Kit according to the manufacturer's instructions (Thermo Fisher No. AM1338). *In vitro* splicing reactions were assembled in 25 μl volumes, each containing 10 μl of splicing extract, 5 μl of 5x splicing buffer (12.5 mM $MgCl_2$, 15% PEG 8000, 300 mM KH_2PO_4 at pH 7.2, and 30 μl nuclease-free water), 5 μl pre-mRNA template from *in vitro* transcription, 0.5 μl 200 mM ATP, 1 μl RNaseOUT Recombinant RNase Inhibitor (Life Technologies No. 10777–019), and 3.5 μl nuclease-free water. For reactions involving the addition of regulatory element RNA, this RNA was added to reactions in lieu of water to the final concentrations indicated in the figures. Reactions

were incubated at 25°C for 40 minutes and then stopped by the addition of 75 μl nuclease-free water and 100 μl of RNA buffer followed immediately by phenol-chloroform RNA extraction. The extracted RNA was DNase-treated and then used for RT-PCR analysis (see below). Equal amounts of DNase-treated RNA were used for all samples for the cDNA synthesis reactions.

Northern blotting

Total RNA aliquots (5–10 μg) were combined with glyoxal buffer to a volume/volume ratio of 5 parts glyoxal buffer to 1 part RNA solution. Glyoxal buffer consists of 60% Dimethyl sulfoxide (Sigma No. D8418), 20% deionized glyoxal (Sigma No. 50649), 1X BPTE [10 mM PIPES free acid (Sigma No. P6757) 30 mM Bis-Tris (Sigma No. B9754), 1 mM EDTA], 4.8% glycerol, and 40 μg/ml ethidium bromide). Glyoxylated RNA was heated for 1 hour at 55°C then cooled in an ice water bath for 10 min. Samples were then vortexed and centrifuged before loading onto agarose gels (between 1.5–2.2% agarose). Samples were separated via electrophoresis at 120V for 2–3 hours in 1X BPTE buffer with constant agitation by magnetic stir bars. Following separation, gels were photographed using a UV imager. Gels were then rinsed in deionized water for 10 minutes followed by a 20 minute equilibration in 75 mM NaOH and two additional 10 minute washes in deionized water. RNA was then blotted onto Amersham Hybond N^+ membranes (VWR No. 95038) overnight with 10X SSPE used as the conducting buffer. 20x SSPE is 25 mM EDTA, 3 M NaCl, and 200 mM NaH_2PO_4-H_2O. RNA was permanently crosslinked to the membrane using UV Stratalinker and hydrated with 2X SSPE. Membranes were pre-hybridized for 30–60 min at 65°C in 10–25 ml of Church's buffer (1% w/v bovine serum albumin, 1 mM EDTA, 0.5 M $NaPO_4$ pH 7.2, 7% SDS), after which 1–100 pmol of radiolabeled riboprobe was added to the buffer and hybridized for at least 8 hours (see below). Following hybridization, membranes were first washed at 65°C for 20 min with 2X SSPE, 0.1% SDS, followed by additional 20 minute washes at 65°C with 1X SSPE, 0.1% SDS, and 0.1X SSPE, 0.1% SDS. Depending on strength of signal following the washes, membranes were exposed in a Kodak K-screen between 15 minutes and 4 days and scanned with a Bio-Rad FX imager. Where indicated, band intensities were quantified using Quantity One volume analysis (Bio-Rad).

DNase treatment, cDNA synthesis and RT-PCR analysis

DNase treatment was performed with Ambion Turbo DNase (Life Technologies No. AM2238). For each sample 40 μg of total RNA was aliquoted to 1.5 ml Eppendorf tubes and combined with 20 μl of 10x DNase Buffer, 4 μl of Turbo DNase, and nuclease-free to a final volume of 200 μl. Mixtures were incubated at 37°C to digest genomic DNA. Samples were then subject to RNA extraction as described above. For cDNA synthesis, 1–5 μg of DNase-treated RNA was combined with 0.4 μl dNTP mix (25 mM), 1 μl 50 ng/μl Random Hexamer (Life Technologies No. N8080127), and nuclease-free water to a final volume of 12 μl. Samples were then heated to 65°C for 5 minutes and chilled on ice for 2 minutes to anneal hexamers. Next, 4 μl 5x First Strand Buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM Magnesium Chloride), 1 μl RNaseOUT, and 1 μl M-MLV Reverse Transcriptase (Life Technologies No. 28025–021) were added to each sample. Samples were mixed gently and incubated at 25°C for 10 minutes, followed by a 50 minute incubation at 42°C and then a 5 minute incubation at 85°C to heat kill the enzymes. Between 1–1.5 μl of cDNA was used for 25 μl PCR reactions with Cy3-labeled primers under standard procedures (see S1 Table for list of primers). Prior to analysis, PCR reactions were mixed with equal volumes of denaturing loading dye (0.02% bromophenol blue, 20 mM EDTA in formamide solvent) and heated at 95°C for 10 minutes. Samples were then analyzed in a Sequi-Gen GT Nucleic Acid Electrophoresis Cell (Bio-Rad No. 165–3860) using 5% denaturing acrylamide gels [8 M urea, 5% Acrylamide:Bis-acrylamide 19:1

(Fisher No. BP1406-1), polymerized with 8.3% ammonium persulfate and 1.6% TEMED]. Samples were separated at 120W for 1–2 hours and gels were dried and analyzed in a Bio-Rad FX imager.

RNA-Immunoprecipitation (RIP)

Prior to RIP, rabbit anti-Myc antibodies (Santa Cruz Biotechnology No. sc-789) were conjugated to Protein G Sepharose beads (GE Healthcare No. 17-0618-01). Briefly, 20 μl of beads per sample were aliquoted into a 1.5 ml Eppendorf tube and centrifuged for 15 seconds at 10000 rpm. The supernatant was removed and beads were washed twice with 500 μl of cold NET-2 buffer (40 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% IGEPAL). 2 μg of antibody per sample was added to the beads and conjugated by gentle rotisserie rotation at 4°C for 1 hour. Beads were then washed three times with 500 μl cold NET-2 buffer to remove unbound antibody.

Yeast cell cultures were grown, harvested, pelleted and flash-frozen as described above, except using low-retention 1.5 ml Eppendorf tubes (Fisher No. 02-681-331). Pellets were resuspended in 600 μl NET-2 buffer and 400 μl acid-washed glass beads. Samples were vortexed at 4°C at maximum intensity for 45 seconds on and 45 seconds off for a total of six cycles. The supernatant was isolated by needle-puncturing the tube and centrifugation into clean 1.5 ml microcentrifuge tubes. Samples were then centrifuged at 4°C for 20 minutes to pellet any remaining cell debris. The resulting supernatant was used to calculate aliquot volumes for immunoprecipitation. Briefly, OD_{260} measurements were made using a Nanodrop 2000 spectrophotometer (Thermo Scientific) to determine $O.D_{260}$ units per ml of sample. Sufficient supernatant aliquots for 2.5 OD_{260} units were added to 20 μl antibody-conjugated Protein G Sepharose beads and brought to a final volume of 400 μl with NET-2 buffer. For the RIP samples, RNA was immunoprecipitated by gentle rotisserie rotation at 4°C for 1 hour. Following IP, samples were washed six times with 1 ml cold NET-2 buffer. RNA was then extracted and prepared for RT-PCR analysis as described above. For the input samples, supernatant aliquots for 2.5 OD_{260} were added to safe-lock Eppendorf tubes. RNA was then extracted and prepared for RT-PCR analysis as described above.

Riboprobe synthesis

Riboprobes were prepared in a 20 μl reaction using Promega T3 RNA Polymerase (Promega No. P4024) according the manufacturer's protocol utilizing DNA templates containing T3 promoter sequences. See S1 Table for list of oligonucleotide primers used to generate T3 templates. Following riboprobe synthesis, the reaction mixture was brought to a total volume of 200 μl with nuclease-free water. The riboprobe solution was then added to the pre-hybridized membranes as described above.

Oligoprobe radiolabeling

Oligonucleotides were labeled with γ-P^{32} ATP using T4 polynucleotide kinase (New England Biolabs No. M0201). Reactions were prepared to a 10 μL volume consisting of 3 μL deionized nuclease-free water, 1 μL 10x polynucleotide kinase buffer (NEB No. B0201), 2 μL 10 mM oligonucleotide solution, 3 μL γ-P^{32} ATP at 10 μCi/μl, and 1μl T4 polynucleotide kinase enzyme. Reactions were incubated at 37°C for 45 minutes and then brought to a final volume of 30 μl with deionized nuclease-free water. The oligonucleotide solution was then added to the pre-hybridized membrane and hybridized at 42°C for at least 8 hours. Washing steps were similar to those described in Northern Blotting procedures except that they were conducted at 42°C and excluded a wash with 0.1x SSPE, 0.1% SDS.

Cloning and bacterial transformation

To clone the *RPL22B* intron into pUG35, the intron was amplified by PCR from genomic DNA using primers that insert BamHI (NEB No. R3136) and EcoRI (NEB No. R0101) restriction sites into the ends of the PCR product. PCR reactions were performed using Phusion Hi-Fi Polymerase (NEB No. M0530) according to the manufacturer's protocol. PCR product was purified using a microcentrifuge spin column according to the manufacturer's protocol (Bio-Pioneer No. PPP-100). Amplification product and pUG35 vector were digested using BamHI and EcoRI restriction enzymes for 2–16 hours. Restriction digest was done using 1 μl 100X BSA, 10 μl NEBuffer, 1 μl CIP (NEB No. M0290, for vector only) and water. Ligation of insert to vector was then performed using T4 DNA Ligase (Life Technologies No. 15224) according the manufacturer's protocol. Plasmids were transformed into DH5α competent E.coli cells (see below) and plated on LB Agar plates with 100 μg/ml ampicillin. Screening of colonies for positive clones was done by growing liquid cultures for miniprep (BioPioneer No. CMIP-100) followed by PCR (see S1 Table for primers used) and positive clones were confirmed by sequencing (Laragen Inc.).

To clone intronless *RPL22A* cDNA into pUG35, total RNA from wild-type cells was reverse-transcribed as described above (see cDNA synthesis). The short length of *RPL22A* exon 1 allowed for the use of a forward primer specific for the spliced cDNA. Forward and reverse primers inserted the BamHI and EcoRI restriction sites into the ends of the PCR product, which was then purified and cloned into pUG35 as described above. Cloning of intronless *RPL22A* cDNA into YEp24 was performed in a similar manner except the reverse primer incorporated the EagI restriction site to the end of the PCR product, which was then purified and cloned into YEp24 as described above.

Site-directed mutagenesis of plasmid DNA

Site-directed mutagenesis was performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies No. 210518). Each reaction combined 5 μl of 10X reaction buffer, 10–100 ng of dsDNA template, 125 ng of oligonucleotide primer #1, 125 ng of oligonucleotide primer #2, 1 μl of dNTP mix, 1.5 μl of QuikSolution reagent, 1 μl of QuikChange Lightning Enzyme and nuclease-free water to a final volume of 50 μl. Oligonucleotides were designed using the QuikChange Primer Design online tool. See S1 Table for a complete list of oligonucleotides used for mutagenesis. Mutagenesis reactions were performed in a thermocycler according to the manufacturer's instructions. After the reaction cycle, each amplification reaction was mixed with 2 μl of DpnI restriction enzyme and then incubated at 37°C for 5 minutes to digest the parental dsDNA. DpnI-treated DNA was then transformed into XL10-Gold ultracompetent cells per the manufacturer's instructions which were then plated onto LB + Ampicillin plates and incubated at 37°C for 16 hours.

DNA transformations into bacteria and yeast

Aliquots of DH5α competent E.coli were added to 1.5 ml Eppendorf tubes and plasmids were added in a 10% v/v amount. Tubes were then incubated on ice for 10 minutes, subject to heat shock for 30 sec at 42°C, and then incubated on ice for 5 minutes. The transformation mixture was resuspended in 1 ml of SOC media (20 mM $MgSO_4$, 2.5 mM KCl, 10 mM NaCl, 2% bacto tryptone, 0.5% yeast extract, 20 mM glucose) and incubated for 1 hour at 37°C. Tubes were then centrifuged for 1 min at 10krpm. After removing most of the supernatant, cell pellets were resuspended in the remaining supernatant (~100 μl) and plated onto LB Agar plates with 100 μg/ml ampicillin. Glass beads were then added to spread cells onto plates until all liquid was absorbed by the plate media. Plated E. coli were then incubated at 37°C. Yeast plasmid and

linear DNA transformations were conducted using the LiAc/SS carrier DNA/PEG method [58].

Spot dilution assays

Agar plates for the spot dilution assays were prepared similarly to the–URA and–HIS-MET liquid media described above, except with the addition of 2% w/v agar. For the testing of cell growth in MMS, media was cooled to ~65°C following sterilization after which 100% MMS (Acros Organics No. 200001–324) was added to the media to a final concentration of 0.005%, 0.01%, or 0.03%.

For the antibiotic screen media was cooled to ~65°C following sterilization after which antibiotics were added to the following concentrations: 0.5 μg/ml anisomycin, 200 μg/ml paromomycin, 100 μg/ml puromycin, 4 mM neomycin, and 2 μg/ml verrucarin A. After the antibiotics or MMS were fully dissolved and mixed, the media plates were poured under sterile conditions and allowed to solidify overnight. Plates were used within two days after preparation. Neomycin, anisomycin, puromycin and verrucarin A were generously provided by Q. Al-Hadid and S. Clarke (UCLA). Paromomycin sulfate was purchased from VWR (Cat. #AAJ61274-06).

Yeast cell cultures were grown in–URA liquid media to exponential phase until an OD_{600} of ~0.4–0.6 was attained. 1 ml volumes of cell culture were then transferred to 1.5 ml Eppendorf tubes and cells were collected by spinning in a microcentrifuge for 30 seconds at 13200 rpm. Cell pellets were resuspended in 1 ml of sterile water and further diluted to $OD_{600} = 0.10$ in a fresh 1.5 ml Eppendorf tube with sterile water. Cells were then subject to five 5-fold series dilutions using sterile water in clean 1.5 ml Eppendorf tubes. For each dilution 4 μl of cell suspension was spotted onto the media plates presented in S10 and S11 Figs. Plates were incubated at 30°C for 1.5–3 days.

Western blot analysis

Cell cultures were grown to exponential phase in rich media (YPD, see above). Cultures were then harvested in 50 ml volumes and centrifuged for 3 minutes at 3000 rpm (Sigma Rotor 11030), washed with deionized water and transferred to 1.5 ml low-retention Eppendorf tubes (Fisher No. 02-681-331). Cells were then pelleted by centrifugation for 30 seconds at 13200 rpm in a benchtop microcentrifuge. Following removal of the supernatant, cells were flash frozen in liquid nitrogen. To lyse the cell pellets, 200 μl of high-salt lysis buffer was added to each sample [200 mM Tris-HCl pH 8.0, 320 mM Ammonium Sulfate, 5 mM $MgCl_2$, 10 mM EGTA pH 8.0, 20 mM EDTA pH 8.0, 1 mM DTT, 20% glycerol, 1 mM PMSF, 20 mM Benzamidine HCl, and 1x Protease Inhibitor cocktail] along with 200 μl acid-washed glass beads. Samples were vortexed at maximum intensity for 8 minutes at 4°C. Cellular debris was pelleted by centrifugation at 3000 rpm for 3 minutes and the supernatant was transferred to clean 1.5 ml Eppendorf tubes. Samples were then centrifuged again at 15000 rpm for 10 minutes at 4°C and the supernatant was transferred to clean 1.5 ml Eppendorf tubes. Lysate protein concentration was quantified with Bio-Rad Bradford protein assay (Bio-Rad No. 500–0001).

To prepare protein electrophoresis samples, 10 μg of protein and water were combined to a total volume of 15 μl, which was then brought to a total volume of 20 μl with the addition of 5 μl 20% β-mercaptoethanol in 4x loading dye. Samples were boiled for 5 minutes and then loaded into 10% SDS-PAGE gels (10% resolving gel composed of 375 mM Tris-HCl pH 8.8, 8% acrylamide, 0.1% SDS, 0.1% APS, and 0.1% TEMED, and 3% stacking gel composed of 126 mM Tris-HCl pH 6.8, 3% acrylamide, 0.1% SDS, 0.1% APS, and 0.1% TEMED). Samples were run for ~1 hour at 180 V. Following electrophoresis, protein was transferred to a 0.22 μm PVDF membrane in 1x Transfer Buffer (20% methanol, 25 mM Tris Base, 192 mM glycine) for

1 hour at 100 V. Protein transfer was confirmed by Coomassie blue staining. Following de-stain, the membrane was incubated in 5% blocking buffer [5% w/v powdered milk in 1x PBS-T (10 mM Sodium Phosphate, 150 mM NaCl, 0.05% Tween 20)] for 1.5 hours at room temperature. Following blocking, the membrane was washed once with 1x PBS-T for 10 minutes. Next, the membrane was incubated with the primary rabbit anti-Myc antibody (Santa Cruz Biotechnology No. sc-789) at a 1:5000 dilution in 5% blocking buffer overnight at 4°C. The membrane was then washed three times for 10 minutes each in 1x PBS-T, and then incubated with the secondary goat anti-rabbit antibody (Pierce No. 31460) at a 1:10000 dilution for one hour at room temperature. The membrane was washed four times for 15 minutes each in 1x PBS-T. To visualize the bands, the membrane was incubated with SuperSignal West Femto Chemiluminescent Substrate (Pierce No. 34096) for 5 minutes and then exposed to x-ray film.

Chromatin immunoprecipitation (ChIP)

Yeast cells expressing C-terminal HA-tagged Prp42p and carrying YEp24 control of *RPL22A* overexpression plasmids were grown in 150 mL cultures to exponential phase at 30°C as described above, after which 37% formaldehyde was added to a final concentration of 1% to initiate crosslinking. Cultures were then incubated with shaking at room temperature for an additional 15 minutes, followed by the addition of glycine to a final concentration of 125 mM and another 5 minute incubation at room temperature to quench the crosslinking. Cells were then pelleted by centrifugation as described above, washed twice in 25 mL of ice cold 1x PBS buffer, collected and flash frozen in liquid nitrogen. The cell pellets were then thawed in 1 mL Lysis Buffer A (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1x protease inhibitor cocktail, 1 mM PMSF) and lysed with the addition of 300 µL acid-washed glass beads and vortexing for 40 minutes at 4°C. The supernatant was then collected into clean 1.5 mL microcentrifuge tubes. Next, the chromatin was subject to shearing by sonication at maximum intensity at 4°C for 7 cycles of 30 seconds on and 1 minute off. The sonicated chromatin was centrifuged for 10 minutes at 4000 RPM and the supernatant was transferred to a fresh 2 mL microcentrifuge tube. 20 µL of cleared chromatin was set aside and flash frozen for the input sample, and 700 µL of chromatin was transferred to a 1.5 mL microcentrifuge tube and incubated at 4°C with 2 µg of rabbit anti-HA primary antibody (Santa Cruz Biotechnology No. Y-11) for 4 hours with gentle rotation. Next, 40 µL of 50% slurry of ImmunoPure Protein A beads (Pierce No. 20333) were added to the chromatin/antibody mixture and the samples were incubated overnight at 4°C with gentle rotation. Following incubation, the samples were briefly spun to pellet the beads and the supernatant was removed. The beads were then subject to the following washes: three washes with 500 µL Lysis Buffer A, one wash with 500 µL Lysis Buffer B (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1x protease inhibitor cocktail, 1 mM PMSF), one wash with 500 µL lithium chloride buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 5 mM EDTA, 1% Triton X-100, 0.5% NP-40, 1x protease inhibitor cocktail, 1 mM PMSF), and one wash with 1 mL 1x TE buffer (100 mM Tris-HCl pH 8, 10 mM EDTA). The supernatant was removed and 250 µL of elution buffer (50 mM Tris-HCl pH 8, 5 mM EDTA, 1% SDS) was added to the beads. 200 µL elution buffer was also added to the input samples that had been thawed on ice. The input and ChIP samples were incubated overnight at 65°C to reverse the crosslinks and then spun briefly to pellet the beads. The supernatant was transferred to fresh 1.5 mL microcentrifuge tubes and treated with 4 µL of 10 mg/mL proteinase K for 1 hour at 42°C. The sample was then cleaned using the QIAquick PCR Purification Kit (Qiagen No. 28104) and eluted into 50 µL of sterile water. Finally, the eluted chromatin was treated with 1 µL of 10 mg/mL RNase A at 37°C for 1 hour and used for PCR analysis.

Supporting Information

S1 Fig. A. *Left*: Comparative table for *RPL22A* and *RPL22B*. Assembled ribosome proportions and percent nucleotide and amino acid identities are provided. *Right*: Scale diagrams comparing *RPL22A* (upper) and *RPL22B* (lower) pre-mRNAs, with the exonic sequences in grey and the introns represented by black lines. The nucleotide positions are labeled with respect to the first nucleotide of the coding sequences. The 5'UTR (blue), alternative 5'-splice site, regulatory element, and RT-PCR primer positions are shown for *RPL22B*. **B.** Northern blot detecting the splicing of *RPL22B* pre-mRNA in wild-type and *upf1*Δ mutant strains carrying the empty pUG23 vector or the *RPL22B* overexpression plasmid. Labeled bands show the unspliced (US) and spliced (S) species. Transcripts were detected using an *RPL22B* 5'UTR riboprobe. *SCR1* was used as a loading control. **C.** Northern blot detecting the splicing of *RPL22A* pre-mRNA in wild-type and *upf1*Δ mutant strains carrying the empty pUG23 vector or the *RPL22A* overexpression plasmid. Transcripts are labeled similarly to panel A and were detected using an *RPL22A* 3'UTR riboprobe. *SCR1* was used as a loading control. **D.** Northern blot detecting the splicing of *RPL22A* pre-mRNA in wild-type and *upf1*Δ mutant strains carrying the empty pUG23 vector or the *RPL22B* overexpression plasmid. Detection methods and labels are similar to panel C. **E.** Northern blot detecting the splicing of *RPL22A* pre-mRNA in wild-type, *upf1*Δ, *rpl22b*Δ, and *upf1*Δ*rpl22b*Δ mutant strains. Detection methods and labels are similar to panel C.
(TIF)

S2 Fig. Systematic search for the *RPL22B* intron regulatory element. A. Northern blot detecting splicing of the *RPL22B* intron reporter transcript expressed from constructs with the full intron or deletion of intronic bases 7 though 152 (Δ7–152) or 153 through 297 (Δ153–297) in strains with or without the chromosomal deletion of *RPL22A*. Labeled bands show the unspliced (US), alternatively-spliced (*AS 5'), and spliced (S) species. Transcripts were detected using a *GFP* ORF riboprobe. *SCR1* was used as a loading control. **B.** RT-PCR analysis of the *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ7–152, or Δ153–297 in strains with or without the chromosomal deletion of *RPL22A*. **C.** Northern blot detecting splicing of the *RPL22B* intron reporter transcript expressed from constructs with the full intron, Δ7–152, or Δ153–297 in strains with or without the chromosomal deletion of *RPL22A* and with or without the deletion of the *RPL22B* alternative 5' splice site. Bands are labeled similarly to panel B. Ethidium bromide-stained 18S rRNA is shown as a loading control.
(TIF)

S3 Fig. Conservation analysis of the *RPL22B* intron. A. The full regulatory element as predicted by Mfold with key nucleotides labeled for reference. Highly conserved nucleotides include both the upstream and downstream nucleotides and are denoted with the black bar. Nucleotides G-153, U-183, A-220 and G-246 have been labeled in colors corresponding to the boxes in the accompanying conservation plots in panel B. **B.** UCSC genome browser snapshots showing strong sequence conservation of intronic nucleotides forming the lower regions of the regulatory element among yeast. *Top panel*: Snapshot of the entire intron. *Center panel*: Snapshot showing conservation of the upstream nucleotides constituting the bottom half of the stem loop, corresponding to intronic nucleotides 153 through 183. *Bottom panel*: Snapshot showing conservation of the downstream nucleotides constituting the bottom half of the stem loop, corresponding to intronic nucleotides 220 through 246. Note that these plots provide the sequence of the Watson strand while *RPL22B* is encoded on the Crick strand. **C.** Structure of the regulatory element as predicted by Mfold when intronic bases 226 through 297 are removed. The nucleotides in green have been determined to promote the regulation of splicing (see text and S5 Fig).
(TIF)

S4 Fig. Targeted mutational analysis of the regulatory element. A. Predicted Mfold structure of the distal end of the Full regulatory element showing the upper and lower distal stems, the hairpin loop, and the intervening RNA loop between the two distal stems. **B.** Northern blot detecting the splicing of the *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ153–188, or deletion of intronic bases 181 through 191 (Δ181–191) or 212 through 223 (Δ212–223) in wild-type cells. Bands are labeled similarly to panels D-E. Transcripts were detecting using a GFP ORF riboprobe. *GAPDH* was used as a loading control. **C.** Predicted Mfold structure of the distal end of the modified construct Δ191–211 UUCG showing the ultra-stable tetraloop. **D.** Northern blot detecting the splicing of the *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ181–191, and the Δ191–211 UUCG construct in wild-type cells. Bands are labeled similarly to panels D-F. *SCR1* was used as a loading control. **E.** Predicted Mfold structures of the "Flip" and "Scramble" constructs, respectively showing the inversion and reordering of the nucleotides of the lower distal stem. **F.** Northern blot detecting the splicing of the *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ181–191, Δ191–211 UUCG, or Δ191–211 UUCG with the lower distal stem bases flipped (Δ191–211 UUCG "Flip") or scrambled (Δ191–2211 UUCG "Scramble") in wild-type cells. Bands are labeled similarly to panels B and D. *SCR1* was used as a loading control.
(TIF)

S5 Fig. Additional targeted mutational analysis of the regulatory element. A. Predicted Mfold structures of constructs tested in panels B and C. Nucleotides of interest are colored red. **B.** Northern blot analysis of *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ181–191, Δ191–211 UUCG, Δ191–211 UUCG with the deletion of the RNA internal loop (Δ191–211 UUCG Δinternal loop), deletion of the lower distal stem and intervening internal loop (Δlower distal stem), mutation of the upstream RNA internal loop sequence from CCCU to AAAC (US Internal Loop AAAC), or mutation of the downstream RNA internal loop sequence from UGAA to CAUU (DS Internal Loop CAUU) in wild-type cells. Labeled bands indicate the unspliced (US) and spliced (S) species. Bands were detected using a GFP ORF riboprobe. *SCR1* was used as a loading control. **C.** Northern blot analysis of *RPL22B* intron reporter transcript expressed from constructs with full intron, Δ181–191, US Internal Loop AAAC, DS Internal Loop CAUU, or mutation of the upstream RNA internal loop sequence from CCCU to UUCA (US Internal Loop UUCA). Bands are labeled similarly to panel A. *SCR1* was used as a loading control.
(TIF)

S6 Fig. The deduced regulatory element is necessary but not sufficient to confer inhibition of splicing *in vivo*. A. Comparison of key intronic properties of *RPL22B* and *RPS21A*. B. Northern blots detecting the splicing patterns of *RPS21A* (upper panel) and *RPL22B* (lower panel) in strains with the empty YEp24 vector or the *RPL22A* overexpression plasmid. The strains carry either the wild-type *RPS21A* intron or mutated versions of the intron harboring either the *RPL22B* regulatory element (22B-reg) or the an 87 nucleotide intronic deletion (Δi87). *GAPDH* was used as a loading control. See S1 Text for additional details.
(TIF)

S7 Fig. Full RT-PCR gel images of the regulatory element addition *in vitro* splicing experiment. These are non-attenuated images of the gels that are presented in truncated form in Fig 4E.
(TIF)

S8 Fig. Western blot of tagged protein and full gel images of RIP experiments. A. Western blot confirmation of the Myc epitope tag on Rpl22Ap. The untagged strain was used as a negative control. Strains expressing Rpl22Ap with a single epitope tag and a Myc-tagged Smd2

protein were used as positive controls. Three separate clones of strains expressing 13-Myc tagged Rpl22Ap were each tested in WT background and the *RPL22B* intron regulatory element deletion background (*RPL22Bi regΔ*). The Myc epitope tag was detected using an anti-Myc primary antibody. Coomassie blue stained proteins are shown as a loading control. **B.** Full unattenuated gel images that were presented in truncated form in Fig 3D.
(TIF)

S9 Fig. Rpl22p and rRNA within the native eukaryotic ribosome. A. PyMol 3D views of the yeast Rpl22 protein and the proximal rRNA sequence. Clockwise from upper left: Rpl22p in the context of all large subunit rRNA, frontal view of Rpl22 and proximal rRNA secondary structure, left side view, and right side view. Protein and RNA structures are from [34]. **B.** 2D representation of the rRNA secondary structure from [35].
(TIF)

S10 Fig. Growth of yeast overexpressing *RPL22* paralogs in the presence of ribosomal inhibitors. Growth of WT and *rpl22aΔ* strains carrying an empty YEp24 vector or *RPL22* paralog overexpression plasmids in various ribosome-inhibiting antibiotics. Following 5-fold spot dilution, plates were incubated at 30°C for the number of days indicated.
(TIF)

S11 Fig. Splicing changes for *RPL22B* and growth of yeast overexpressing *RPL22* paralogs in the presence of MMS. A. Quantification of spliced *RPL22B* transcript in WT and *RPL22Bi regΔ* strains after three hours of treatment with 0.05% MMS reported as a percentage of spliced transcript that was present prior to treatment (t = 0). The values shown are mean ± 1 standard deviation measured from Northern blot experiments consisting of three independent biological replicates for each strain. B. Quantification of unspliced and spliced *RPL22B* transcripts in wild-type, *RPL22Bi regΔ*, and *rpl22aΔ* strains at steady state and following three hours of treatment with 0.05% MMS, given as percentage of the total transcript. The values shown are mean ± 1 standard deviation measured from Northern blot experiments consisting of three independent biological replicates for each strain. C. Wild-type cells carrying an empty YEp24 vector or plasmids overexpressing the intronless cDNA of *RPL22A* and *RPL22B* were tested for growth in 0%, 0.005%, 0.01% and 0.03% MMS. Following 5-fold spot dilution, plates were incubated at 30°C for the number of days indicated.
(TIF)

S12 Fig. rRNA transcription and maturation is inhibited by cadmium chloride treatment. Northern blots detecting the 35S rRNA and other rRNA processing intermediates. Cells were grown in YPD containing 50 μM $CdCl_2$ for 120 minutes. The 35S rRNA and processing intermediates were detected using oligoprobes complementary to transcripts arising from the *RDN1* locus as shown in the lower diagram. Sequences for O1660 and O1663 were obtained from Lindahl et al. [59]. *SCR1* was used as a loading control.
(TIF)

S13 Fig. Full gel image of the *in vitro* splicing and RT-PCR experiment utilizing wild-type splicing extracts. This is a non-attenuated image of the gel that is presented in truncated form in Fig 4C.
(TIF)

S14 Fig. Full gel image of the *in vitro* splicing and RT-PCR experiment utilizing wild-type and *RPL22A* overexpression splicing extracts. This is a non-attenuated image of the gel that is presented in truncated form in Fig 4D.
(TIF)

S1 Text. Comprehensive description of experiments to elucidate the regulatory element.
(DOCX)

S1 Table. List of yeast strains, plasmids, and oligonucleotides used in this study.
(XLSX)

Acknowledgments

We thank Kevin Roy and Jonathan Warner for thoughtful discussion and critical reading of the manuscript. We thank Tracy Johnson and Brian Kennedy for sharing of materials and Lauren Neves for helpful advice. We thank Jocelyn Almanza for extensive technical assistance. We also wish to acknowledge an anonymous reviewer for insightful experimental suggestions.

Author Contributions

Conceived and designed the experiments: JG GC. Performed the experiments: JG. Analyzed the data: JG GC. Wrote the paper: JG GC.

References

1. Jelinsky SA, Samson LD (1999) Global response of Saccharomyces cerevisiae to an alkylating agent. Proc Natl Acad Sci U S A 96: 1486–1491. PMID: 9990050
2. Hu H, Li X (2007) Transcriptional regulation in eukaryotic ribosomal protein genes. Genomics 90: 421–423. PMID: 17707610
3. Parenteau J, Durand M, Veronneau S, Lacombe AA, Morin G, et al. (2008) Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol Biol Cell 19: 1932–1941. doi: 10.1091/mbc.E07-12-1254 PMID: 18287520
4. Parenteau J, Durand M, Morin G, Gagnon J, Lucier JF, et al. (2011) Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. Cell 147: 320–331. doi: 10.1016/j.cell.2011.08.044 PMID: 22000012
5. Libri D, Stutz F, McCarthy T, Rosbash M (1995) RNA structural patterns and splicing: Molecular basis for an RNA-based enhancer. RNA 1: 425–436. PMID: 7493320
6. Buratti E, Baralle FE (2004) Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. Mol and Cell Bio 24: 2055–2066.
7. Warner JR, McIntosh KB (2009) How Common Are Extraribosomal Functions of Ribosomal Proteins? Molecular Cell 34: 3–11. doi: 10.1016/j.molcel.2009.03.006 PMID: 19362532
8. Eng FJ, Warner JR (1991) Structural basis for the regulation of splicing of a yeast messenger RNA. Cell 65: 797–804. PMID: 2040015
9. Fewell SW, Woolford JL Jr. (1999) Ribosomal protein S14 of Saccharomyces cerevisiae regulates its expression by binding to RPS14B pre-mRNA and to 18S rRNA. Mol Cell Biol 19: 826–834. PMID: 9858605
10. Petibon C, Parenteau J, Catala M, Elela SA (2016) Introns regulate the production of ribosomal proteins by modulating splicing of duplicated ribosomal protein genes. Nucleic Acids Res.
11. He F, Li C, Roy B, Jacobson A (2014) Yeast Edc3 targets RPS28B mRNA for decapping by binding to a 3' untranslated region decay-inducing regulatory element. Mol Cell Biol 34: 1438–1451. doi: 10.1128/MCB.01584-13 PMID: 24492965
12. Gudipati RK, Neil H, Feuerbach F, Malabat C, Jacquier A (2012) The yeast RPL9B gene is regulated by modulation between two modes of transcription termination. EMBO J 31: 2427–2437. doi: 10.1038/emboj.2012.81 PMID: 22505027
13. O'Leary MN, Schreiber KH, Zhang Y, Duc AC, Rao S, et al. (2013) The ribosomal protein Rpl22 controls ribosome composition by directly repressing expression of its own paralog, Rpl22l1. PLoS Genet 9: e1003708. doi: 10.1371/journal.pgen.1003708 PMID: 23990801
14. Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF (2014) Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae. PLoS Genet 10: e1004249. doi: 10.1371/journal.pgen.1004249 PMID: 24722551

15. DeLuna A, Springer M, Kirschner MW, Kishony R (2010) Need-based up-regulation of protein levels in response to deletion of their duplicate genes. PLoS Biol 8: e1000347. doi: 10.1371/journal.pbio.1000347 PMID: 20361019

16. Hughes TR, Roberts CJ, Dai H, Jones AR, Meyer MR, et al. (2000) Widespread aneuploidy revealed by DNA microarray expression profiling. Nat Genet 25: 333–337. PMID: 10888885

17. Steffen KK, McCormick MA, Pham KM, MacKay VL, Delaney JR, et al. (2012) Ribosome deficiency protects against ER stress in Saccharomyces cerevisiae. Genetics 191: 107–118. doi: 10.1534/genetics.111.136549 PMID: 22377630

18. Chan CT, Pang YL, Deng W, Babu IR, Dyavaiah M, et al. (2012) Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. Nat Commun 3: 937. doi: 10.1038/ncomms1938 PMID: 22760636

19. Lu C, Brauer MJ, Botstein D (2009) Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell 20: 891–903. doi: 10.1091/mbc.E08-08-0852 PMID: 19056679

20. Steffen KK, MacKay VL, Kerr EO, Tsuchiya M, Hu D, et al. (2008) Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. Cell 133: 292–302. doi: 10.1016/j.cell.2008.02.037 PMID: 18423200

21. Heyer EE, Moore MJ (2016) Redefining the Translational Status of 80S Monosomes. Cell 164: 757–769. doi: 10.1016/j.cell.2016.01.003 PMID: 26871635

22. Mumberg D, Muller R, Funk M (1994) Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res 22: 5767–5768. PMID: 7838736

23. Parker R (2012) RNA Degradation in Saccharomyces cerevisae. Genetics 191: 671–702. doi: 10.1534/genetics.111.137265 PMID: 22785621

24. Anderson JS, Parker R (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. EMBO J 17: 1497–1506. PMID: 9482746

25. Sun M, Schwalb B, Pirkl N, Maier KC, Schenk A, et al. (2013) Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. Mol Cell 52: 52–62. doi: 10.1016/j.molcel.2013.09.010 PMID: 24119399

26. Sayani S, Chanfreau GF (2012) Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in Saccharomyces cerevisiae. RNA 18: 1563–1572. doi: 10.1261/rna.033779.112 PMID: 22753783

27. Bergkessel M, Whitworth GB, Guthrie C (2011) Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA 17: 1461–1478. doi: 10.1261/rna.2754011 PMID: 21697354

28. Bracken AP, Bond U (1999) Reassembly and protection of small nuclear ribonucleoprotein particles by heat shock proteins in yeast cells. RNA 5: 1586–1596. PMID: 10606269

29. Imamachi N, Tani H, Akimitsu N (2012) Up-frameshift protein 1 (UPF1): multitalented entertainer in RNA decay. Drug Discov Ther 6: 55–61. PMID: 22622014

30. Varsally W, Brogna S (2012) UPF1 involvement in nuclear functions. Biochem Soc Trans 40: 778–783. doi: 10.1042/BST20120052 PMID: 22817733

31. Sun M, Schwalb B, Schulz D, Pirkl N, Etzold S, et al. (2012) Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Res 22: 1350–1359. doi: 10.1101/gr.130161.111 PMID: 22466169

32. Storici F, Resnick MA (2006) The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol 409: 329–345. PMID: 16793410

33. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415. PMID: 12824337

34. Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, et al. (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. Science 334: 1524–1529. doi: 10.1126/science.1212642 PMID: 22096102

35. Petrov AS, Bernier CR, Gulen B, Waterbury CC, Hershkovits E, et al. (2014) Secondary structures of rRNAs from all three domains of life. PLoS One 9: e88222. doi: 10.1371/journal.pone.0088222 PMID: 24505437

36. Dobbelstein M, Shenk T (1995) In vitro selection of RNA ligands for the ribosomal L22 protein associated with Epstein-Barr virus-expressed RNA by using randomized and cDNA-derived RNA libraries. J Virol 69: 8027–8034. PMID: 7494316

37. Haruki H, Nishikawa J, Laemmli UK (2008) The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol Cell 31: 925–932. doi: 10.1016/j.molcel.2008.07.020 PMID: 18922474

38. Zemp I, Kutay U (2007) Nuclear export and cytoplasmic maturation of ribosomal subunits. FEBS Lett 581: 2783–2793. PMID: 17509569

39. Kotovic KM, Lockshon D, Boric L, Neugebauer KM (2003) Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. Mol Cell Biol 23: 5768–5779. PMID: 12897147

40. Al-Hadid Q, Roy K, Munroe W, Dzialo MC, Chanfreau GF, et al. (2014) Histidine methylation of yeast ribosomal protein Rpl3p is required for proper 60S subunit assembly. Mol Cell Biol 34: 2903–2916. doi: 10.1128/MCB.01634-13 PMID: 24865971

41. Esguerra J, Warringer J, Blomberg A (2008) Functional importance of individual rRNA 2'-O-ribose methylations revealed by high-resolution phenotyping. RNA 14: 649–656. doi: 10.1261/rna.845808 PMID: 18256246

42. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C (2007) Rapid, transcript-specific changes in splicing in response to environmental stress. Mol Cell 27: 928–937. PMID: 17889666

43. Urban J, Soulard A, Huber A, Lippman S, Mukhopadhyay D, et al. (2007) Sch9 is a major target of TORC1 in Saccharomyces cerevisiae. Mol Cell 26: 663–674. PMID: 17560372

44. Zhou L, Le Roux G, Ducrot C, Chedin S, Labarre J, et al. (2013) Repression of class I transcription by cadmium is mediated by the protein phosphatase 2A. Nucleic Acids Res 41: 6087–6097. doi: 10.1093/nar/gkt335 PMID: 23640330

45. Jenner L, Melnikov S, Garreau de Loubresse N, Ben-Shem A, Iskakova M, et al. (2012) Crystal structure of the 80S yeast ribosome. Curr Opin Struct Biol 22: 759–767. doi: 10.1016/j.sbi.2012.07.013 PMID: 22884264

46. Tardiff DF, Lacadie SA, Rosbash M (2006) A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. Mol Cell 24: 917–929. PMID: 17189193

47. Chao JA, Williamson JR (2004) Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. Structure 12: 1165–1176. PMID: 15242593

48. Vilardell J, Warner JR (1993) Regulation of splicing at an intermediate step in the formation of the spliceosome. Genes Dev 8: 211–220.

49. Meyer M, Plass M, Perez-Valle J, Eyras E, Vilardell J (2011) Deciphering 3'ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. Mol Cell 43: 1033–1039. doi: 10.1016/j.molcel.2011.07.030 PMID: 21925391

50. Antoniali G, Lirussi L, Poletto M, Tell G (2014) Emerging roles of the nucleolus in regulating the DNA damage response: the noncanonical DNA repair enzyme APE1/Ref-1 as a paradigmatical example. Antioxid Redox Signal 20: 621–639. doi: 10.1089/ars.2013.5491 PMID: 23879289

51. Nalabothula N, Indig FE, Carrier F (2010) The Nucleolus Takes Control of Protein Trafficking Under Cellular Stress. Mol Cell Pharmacol 2: 203–212. PMID: 21499571

52. Roy K, Chanfreau G (2014) Stress-induced nuclear RNA degradation pathways regulate yeast bromodomain factor 2 to promote cell survival. PLoS Genet 10: e1004661. doi: 10.1371/journal.pgen.1004661 PMID: 25232960

53. Munding EM, Shiue L, Katzman S, Donohue JP, Ares M (2013) Competition between Pre-mRNAs for the Splicing Machinery Drives Global Regulation of Splicing. Mol Cell 51: 338–348. doi: 10.1016/j.molcel.2013.06.012 PMID: 23891561

54. Fahl SP, Harris B, Coffey F, Wiest DL (2015) Rpl22 Loss Impairs the Development of B Lymphocytes by Activating a p53-Dependent Checkpoint. J Immunol 194: 200–209. PMID: 25416806

55. Stadanlick JE, Zhang Z, Lee SY, Hemann M, Biery M, et al. (2011) Developmental arrest of T cells in Rpl22-deficient mice is dependent upon multiple p53 effectors. J Immunol 187: 664–675. doi: 10.4049/jimmunol.1100029 PMID: 21690328

56. Meek DW (2009) Tumour suppression by p53: a role for the DNA damage response? Nat Rev Cancer 9: 714–723. doi: 10.1038/nrc2716 PMID: 19730431

57. Bader DM, Wilkening S, Lin G, Tekkedil MM, Dietrich K, et al. (2015) Negative feedback buffers effects of regulatory variants. Mol Syst Biol 11: 785. doi: 10.15252/msb.20145844 PMID: 25634765

58. Geitz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2: 31–34. PMID: 17401334

59. Lindahl L, Bommankanti A, Li X, Hayden L, Jones A, et al. (2009) RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. RNA 15: 1407–1416. doi: 10.1261/rna.1302909 PMID: 19465684

Supporting Information

S1 Text

Binary search approach

We began the search for the *RPL22B* intronic regulatory element by making systematic bifurcating deletion of the GFP reporter transcript as outlined in Fig 3A. Our initial experimental system utilized the *rpl22a*Δ strain as a control for uninhibited splicing based on the presumption that the removal of the regulatory element from the intron would phenocopy the loss of splicing inhibition for *RPL22B* that is observed in the absence of *RPL22A* (Fig 1C). Indeed, the splicing of the reporter transcript containing the full intron was greatly enhanced in the *rpl22a*Δ strain compared to wildtype (S2A Fig, lanes 1 and 2). The first set of deletions removed either intronic nucleotides 7 through 152 (Δ7-152) or nucleotides 153 through 297 (Δ153-297) from the reporter (Fig 3A). The former deletion had no impact on reporter splicing efficiency in the presence of *RPL22A*, indicating that this region did not harbor the regulatory sequence (S2A Fig, lanes 3 and 4). Surprisingly, the Δ153-297 deletion also did not enhance the splicing of the reporter (S2A Fig, lane 5). Furthermore, splicing efficiency was not enhanced in the *rpl22a*Δ strain for the Δ153-297 construct (S2A Fig, lane 6), suggesting that the assay was not functioning as expected in this particular deletion construct. However, an additional splice product running slightly higher than the normal spliced product was also detected in this deletion. We therefore re-analyzed the splicing patterns by RT-PCR and we found that the deletion of nucleotides 153 through 297 promoted the use of the alternative 5' splice site (AS 5') previously described for *RPL22B* (S2B Fig), which in turn interfered with the assay. A band representing the alternatively spliced species was clearly visible in the Δ153-297 strains in addition to the unspliced and normal splice products (S2B Fig). The reason for this change in preferred splice site selection in this deletion construct is not clear, but is possibly related to the impact of physical spacing between the splice sites and the branch point [1-3].

Because the presence of the AS 5' within the *RPL22B* intron obstructed our ability to analyze sequences required for *RPL22B* splicing regulation, we removed the AS 5' through site-directed mutagenesis for the remainder of our search for the regulatory element. The deletion of the AS 5' restored the functionality of the assay, revealing a substantial increase in splicing efficiency in the Δ153-297 construct when compared to the full-length construct even in the presence of *RPL22A* (S2C Fig, lanes 3 and 7). Thus, the regulatory element resides within the region spanning nucleotides 153 through 297 of the *RPL22B* intron. Interestingly, deletion of the AS 5' also appeared to reduce overall splicing efficiency at the annotated 5' splice site in the full-length transcript (S2C Fig, lanes 2 and 4), hinting that the presence of dual 5' splice sites may increase spliceosome assembly efficiency *in vivo* for *RPL22B*.

The second set of intronic truncations consisted of deletions of nucleotides 153 through 225 and 226 through 297. The Δ153-225 deletion eliminated the inhibition of splicing while the Δ226-297 deletion maintained it (Fig 3A, top right panel), thus narrowing the regulatory element to the span of nucleotides 153 through 225. Finally, the third series of truncations deleted nucleotides 153 through 188 and 189 through 225. Splicing inhibition was eliminated in both cases (Fig 3A, bottom right panel), suggesting that the regulatory element consists of nucleotides from both of these regions. We were therefore unable to further pinpoint the regulatory element using the binary search approach and were required to conduct more specific targeted mutations of this region.

The intronic segment containing the regulatory element is predicted to form a structured stem loop

Upon identifying the general location of the regulatory element as within the region bordered by intronic nucleotides 153 through 225, we examined the predicted structure of the *RPL22B* intron as determined by the Mfold web server. We found that this span of intronic nucleotides has the potential to fold into a stem loop that includes nucleotides 153 through 246 and that might constitute an RNA secondary structure required for splicing inhibition (S3A Fig). Intriguingly, the nucleotides constituting approximately the lower half of this stem loop are highly conserved amongst related yeast species (S3A Fig and S3B Fig). This contrasts with the general lack of conservation the intronic sequence, suggesting

that these sequences may be functionally relevant and perhaps serve a role in the inhibition of splicing. However, our deletion construct Δ226-297 maintained inhibition of splicing, even though it is predicted to remove most of the lower conserved regions from the secondary structure and leave only the upper portions structurally intact (Fig 3A and S3C Fig). This suggests that the majority of conserved nucleotides are not necessary for splicing inhibition, although we cannot rule out the possibility that they may indirectly contribute to the overall structure by enhancing the exposure of the remainder of the stem loop or contribute to the folding of the tertiary structure.

Secondary structure, not sequence specificity, of the lower distal stem is required for inhibition of splicing

Our previous experiments demonstrated the necessity of the approximate upper (distal) half of the stem loop of this secondary structure for the bestowal of splicing inhibition. This region consists of a lower stem, an intervening internal loop, and an upper stem that is capped by an AAUGC hairpin pentaloop (S4A Fig). All of these features are potential candidates for RNA-protein interactions [4, 5]. To further dissect these components, we disrupted this region by creating three additional independent deletion constructs. The first, Δ181-191, removes the 5' nucleotides from the lower stem and was predicted by Mfold to dramatically reshape the intron such that the putative regulatory element no longer forms. The second construct, Δ192-211, removes the upper stem loop and intervening internal loop while maintaining the lower stem, closing it with a UUACU pentaloop and also preserving the overall secondary structure of the entire intron. The third construct, Δ212-223, removes the 3' nucleotides from the lower stem and is likewise expected to completely alter the intronic secondary structure. Not surprisingly, constructs Δ181-191 and Δ212-223 did not exhibit splicing regulation by Rpl22p (S4B Fig). However, splicing regulation by Rpl22p was maintained in construct Δ192-211, suggesting that the upper stem and hairpin loop are not required for recognition by Rpl22p and subsequent inhibitory effects, while the lower stem is essential (S4B Fig).

We next were interested in determining what facets of the lower stem region contribute to splicing regulation. First, we truncated the full structure by removing the upper stem and its hairpin loop (both of which were demonstrated as nonessential in the previous experiment) and closing the lower stem with a UUCG tetraloop, creating the construct Δ191-211 UUCG (S4C Fig). This hairpin sequence is a member of the UNCG family of thermodynamically stable loop sequences [6], enabling further manipulations of the remaining stem nucleotides with minimal risk of dramatically altering the structure. Consistent with our previous experiments, the replacement of the upper distal stem loop with the UUCG tetraloop sequence retained splicing regulation (S4D Fig). To unambiguously determine whether sequence identity of the lower stem is required to maintain the regulatory mechanism, we generated two additional constructs in which the identities of the stem nucleotides were altered but secondary structure was maintained. In the first construct ("Flip"), we reversed the positions of the upstream 5' and downstream 3' stem sequences (S4E Fig, left). In the second construct ("Scramble"), we shifted the positions of the stem base pairs while maintaining the identities of the pairs themselves – that is, we maintained the two G-C pairs and five A-U pairs (S4E Fig, right). Neither mutation enhanced the splicing of the reporter (S4G Fig), supporting the argument that the structure and not the sequence identity of this region is necessary for the inhibition of splicing. However, although our previous constructs that disrupted this stem (Δ181-191 and Δ212-223) demonstrated a loss of regulation by Rpl22p, we could not rule out that this was a result of the dramatic restructuring of the intron that may also have affected other critical areas of the regulatory element. To account for this possibility, we deleted this lower stem from the full wild-type structure, leaving the remainder of the regulatory element and the rest of the intron intact (S5A Fig, Δlower distal stem structure). This mutation resulted in a loss of splicing regulation (S5B Fig, lane 5), confirming our hypothesis that the lower distal stem structure is necessary for this mechanism to function, though it is flexible in its sequence identity.

The downstream nucleotides of the putative RNA internal loop are likely paired with upstream nucleotides

Truncations of *RPL22B* intron reporter revealed that the upper distal stem loop (nucleotides 191 through 211), the intervening RNA internal loop, and the nucleotides downstream of position 225 are unnecessary for splicing inhibition (Fig 3A). Furthermore, the deletion of nucleotides 226 through 297 eliminated the participation of nucleotides 153 through 172 in the stem (S3C Fig). The resulting preservation of splicing regulation implies that the intronic nucleotides upstream of G153 are not required for the inhibitory mechanism (Fig 3A). Thirdly, our previous experiments showed that the lower distal stem is required for the inhibition of splicing (S4 Fig and S5 Fig). Therefore, the final remaining component in this region that could additionally contribute to the regulatory element was the putative "RNA internal loop" immediately below the lower distal stem formed by upstream nucleotides 178 through 181 and downstream nucleotides 221 through 224 (S5A Fig, see "Full" structure). We deleted these nucleotides from the simplified UUCG tetraloop reporter (S5A Fig, Δ191-211 UUCG Δinternal loop) which resulted in a notable increase in splicing, albeit not to the extent seen when the entire regulatory element is disrupted (S5B Fig, lanes 2 and 4). This suggests that the putative internal loop is partially involved in the inhibition of splicing. In agreement with this notion, the Δlower distal stem construct retains the RNA internal loop nucleotides and yet loses the ability to inhibit splicing (S5A Fig and S5B Fig), suggesting that this loop or its constituent nucleotides contribute to the regulatory mechanism but alone are insufficient to confer it.

We also investigated whether the sequence identity of this potential internal loop region is crucial for its role as part of the regulatory element. For this experiment we worked from the full intron construct (S5A Fig) and created two new reporters. In the first, the four upstream internal loop nucleotides were mutated from CCCU to AAAC (S5A Fig, "US Internal Loop AAAC"). In the second, the four downstream internal loop nucleotides were mutated from UGAA to CAUU (S5A Fig, "DS Internal Loop CAUU"). In both cases the mutations were predicted to maintain the overall structure of the entire stem loop. Surprisingly, the mutation of the downstream internal loop nucleotides appeared to at least partially mitigate splicing inhibition while mutation of the upstream internal loop had no effect (S5B Fig, lanes 6-7). This result is seemingly contradictory to the behavior of the Δ226-297 construct, which eliminated the downstream internal loop and yet did not disrupt the inhibition of splicing (Fig 3A and S3C Fig). Importantly, the downstream internal loop nucleotides UGAA were retained in that construct but were no longer predicted to constitute a single-stranded internal loop. Instead, they were paired with upstream nucleotides 173 through 176 (UUCA), forming a perfect complement (S3C Fig). Based on this evidence, we hypothesized that the inhibition of splicing depends on the nucleotide identity of the downstream UGAA sequence and that it may actually be a paired structure in its most stable form instead of a single-stranded loop region as predicted by Mfold. To test this idea, we manipulated the full intron reporter construct by mutating the upstream internal loop nucleotides to UUCA, thus establishing a perfect complementarity with the downstream nucleotides in the presumed internal loop. The predicted Mfold structure shows that this mutation closes the internal loop completely by pairing each constituent nucleotide (S5A Fig, "US Internal Loop UUCA"). Closing of the internal loop in this manner did not increase splicing efficiency (S5C Fig), suggesting that the downstream RNA internal loop is actually a paired sequence. Based on our experimental evidence, we speculate that in its most stable configuration nucleotides U221 through A224 of the putative RNA internal loop form base pairs with upstream nucleotides U173 through A176. Indeed, constraining this pairing using Mfold results in only one stable predicted structure that very closely resembles the Δ226-297 construct (compare Fig 3B and S3C Fig). Importantly, this experimentally-deduced structure retains the essential lower distal stem.

The deduced regulatory element is necessary but not sufficient for regulatory activity *in vivo*

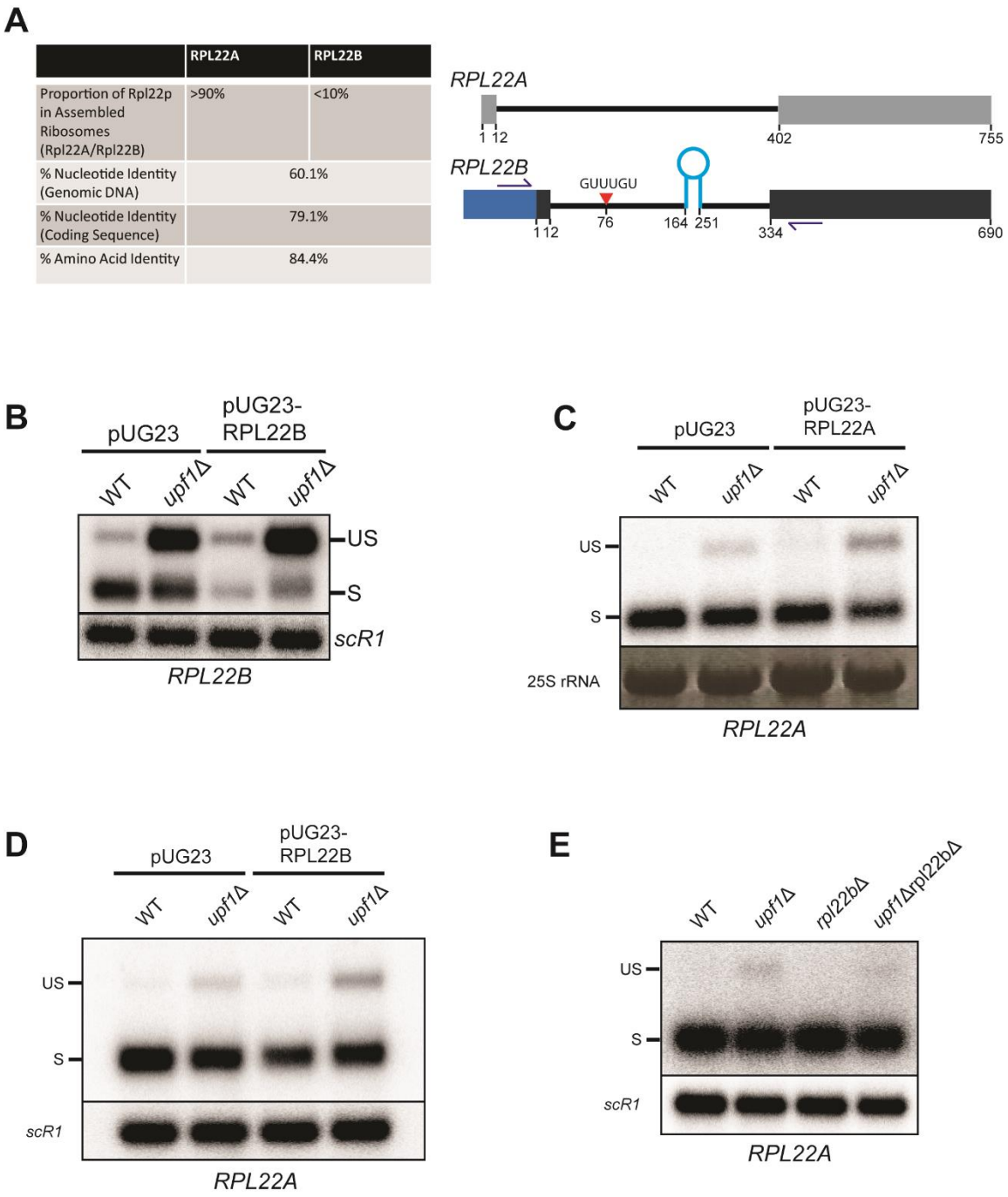
The experiments thus far have indicated that the *RPL22B* regulatory element consisting of intronic nucleotides 153 through 239 is necessary for the inhibition of splicing (Fig 3C). To test whether this region is also sufficient for achieve autoregulation, we transposed this span of nucleotides into the

intron of *RPS21A*, another RPG whose intron is of similar overall length to that of *RPL22B* and whose branch point is located in a similar relative position relative to the 5'-splice site (S6 Fig). To achieve transposition without impacting overall intron size, we first removed 87 nucleotides from the *RPS21A* intron that occupied the same region of that intron as the regulatory element in the *RPL22B* intron. We then inserted the 87 nucleotides of the *RPL22B* regulatory element into this region using the delitto perfetto technique. The *RPL22A* overexpression plasmid was then introduced into this strain and examined the *RPS21A* transcript in wild-type and NMD mutant background strains to test whether the chimeric intron would experience inhibition of splicing. Although both the deletion of the 87 nucleotides from the *RPS21A* intron as well as the transposition of the *RPL22B* intronic nucleotides impacted the susceptibility of the *RPS21A* pre-mRNA to NMD, it did not render the pre-mRNA sensitive to splicing inhibition by Rpl22p (S6 Fig, compare lanes 1-4 to lanes 5-8). Thus, this span of 87 nucleotides is insufficient to confer inhibition of splicing *in vivo*.

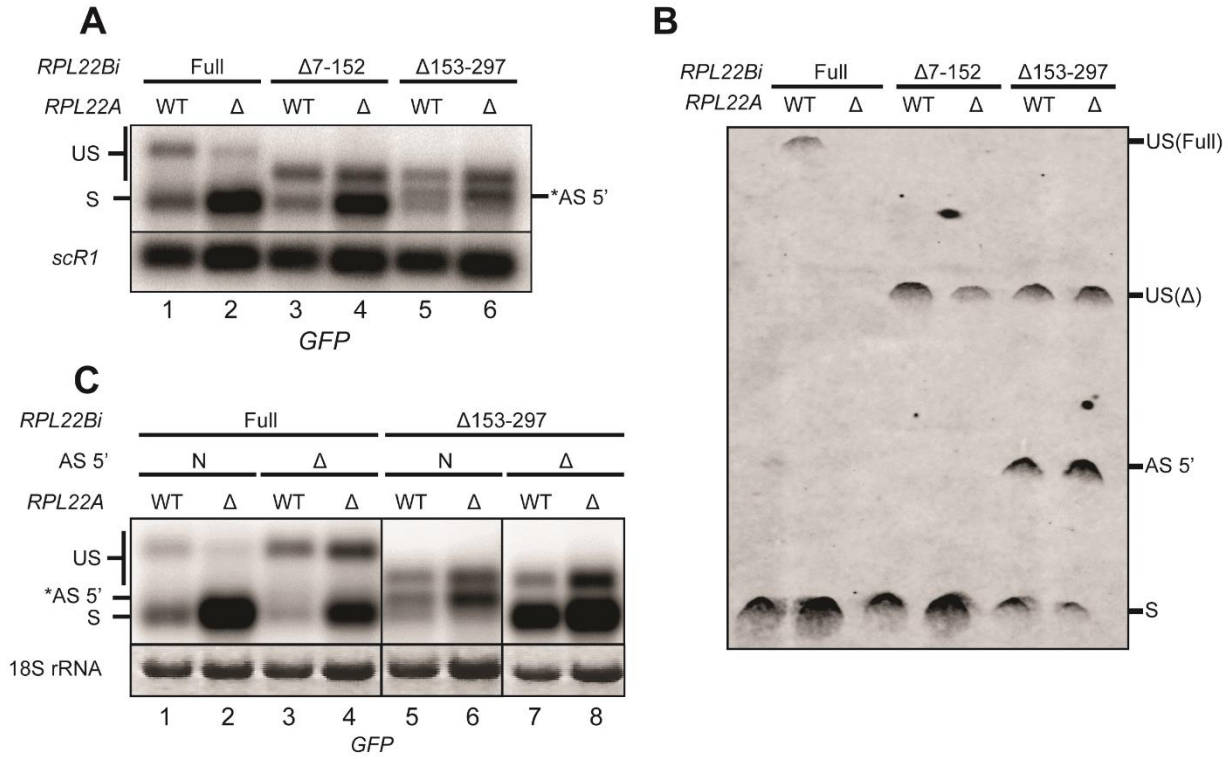
Supporting References

1. Cellini A, Felder E, Rossi JJ. Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3' splice site. EMBO J. 1986 May;5(5):1023-30.
2. Crawford DJ, Hoskins AA, Friedman LJ, Gelles J, Moore MJ. Single-molecule colocalization FRET evidence that spliceosome activation precedes stable approach of 5' splice site and branch site. Proc Natl Acad Sci U S A. 2013 Apr 23;110(17):6783-8.
3. Thompson-Jager S, Domdey H. Yeast pre-mRNA splicing requires a minimum distance between the 5' splice site and the internal branch acceptor site. Mol Cell Biol. 1987 Nov;7(11):4010-6.
4. Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. Protein-RNA interactions: a structural analysis. Nucleic Acids Res. 2001 Feb 15;29(4):943-54.
5. Nagai K. RNA-protein interactions. Current Opinion in Structural Biology. 1992;2(1):131-7.
6. Molinaro M, Tinoco I. Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications. Nucleic Acids Res. 1995;23(15):3056-63.

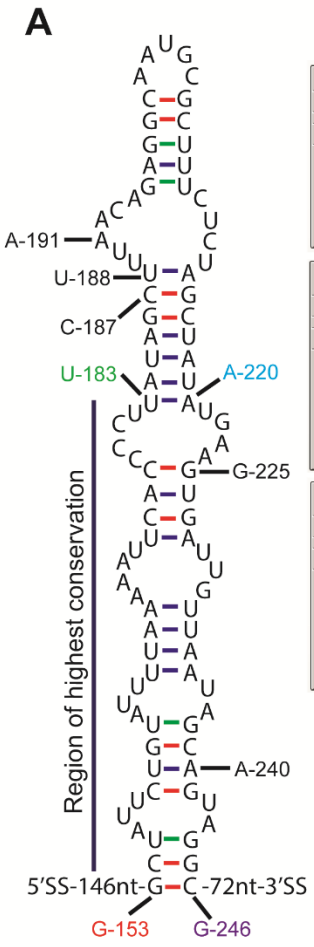
S1 Fig

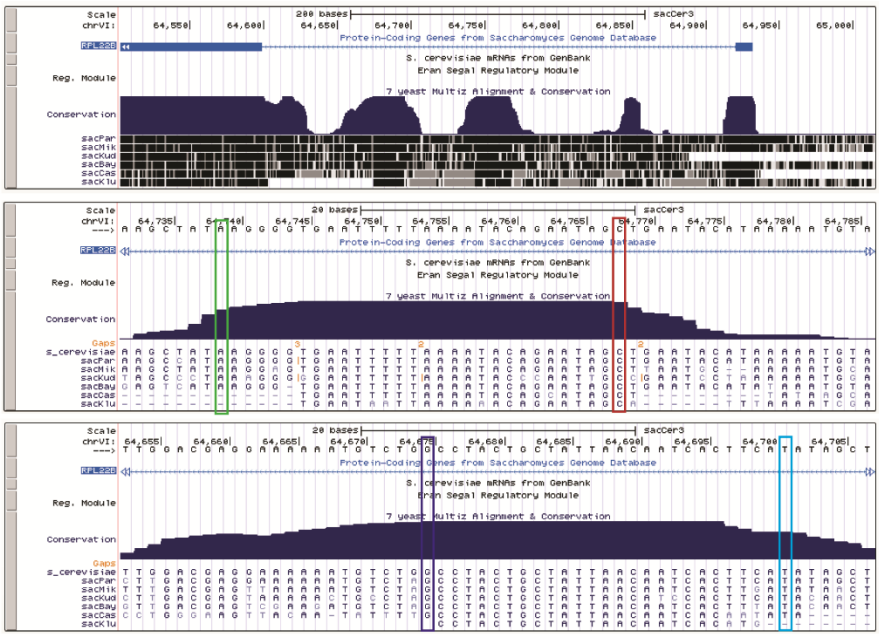

A
RPL22A
RPL22B
Proportion of Rpl22p in Assembled Ribosomes (Rpl22A/Rpl22B)
>90%
<10%
% Nucleotide Identity (Genomic DNA)
60.1%
% Nucleotide Identity (Coding Sequence)
79.1%
% Amino Acid Identity
84.4%
RPL22A
1 12
402
755
RPL22B
GUUUGU
1 12
76
164 251
334
690
B
pUG23
pUG23-RPL22B
WT
upf1Δ
WT
upf1Δ
US
S
scR1
RPL22B
C
pUG23
pUG23-RPL22A
WT
upf1Δ
WT
upf1Δ
US
S
25S rRNA
RPL22A
D
pUG23
pUG23-RPL22B
WT
upf1Δ
WT
upf1Δ
US
S
scR1
RPL22A
E
WT
upf1Δ
rpl22bΔ
upf1Δrpl22bΔ
US
S
scR1
RPL22A

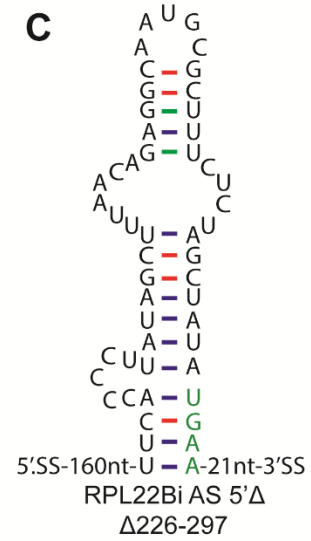
S2 Fig


A
RPL22Bi
Full
Δ7-152
Δ153-297
RPL22A
WT
Δ
WT
Δ
WT
Δ
US
S
scR1
*AS 5'
1
2
3
4
5
6
GFP
B
RPL22Bi
Full
Δ7-152
Δ153-297
RPL22A
WT
Δ
WT
Δ
WT
Δ
US(Full)
US(Δ)
AS 5'
S
C
RPL22Bi
Full
Δ153-297
AS 5'
N
Δ
N
Δ
RPL22A
WT
Δ
WT
Δ
WT
Δ
WT
Δ
US
*AS 5'
S
18S rRNA
1
2
3
4
5
6
7
8
GFP

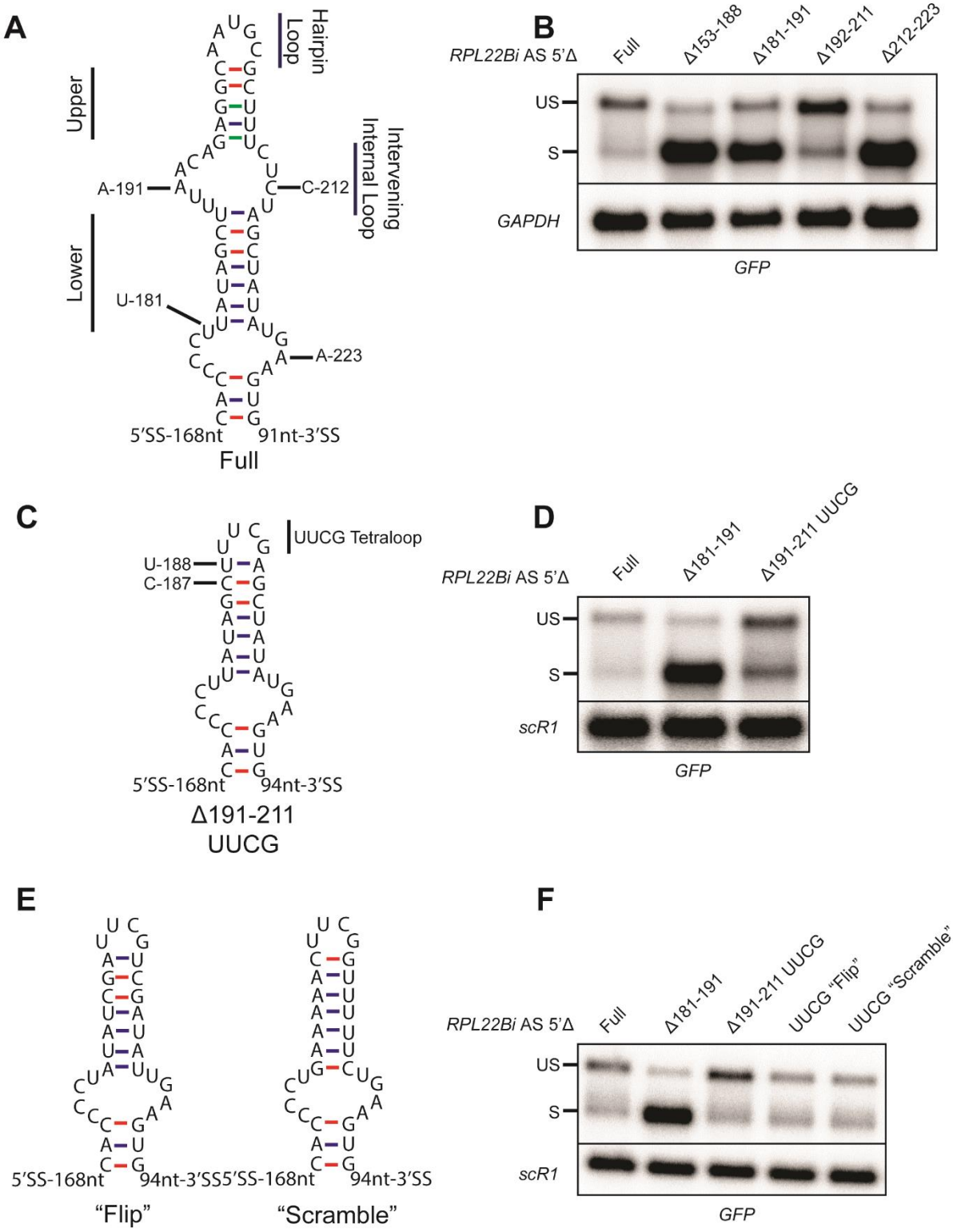
S3 Fig


A
A-191
U-188
C-187
U-183
A-220
G-225
A-240
Region of highest conservation
5′SS-146nt-G
C-72nt-3′SS
G-153
G-246

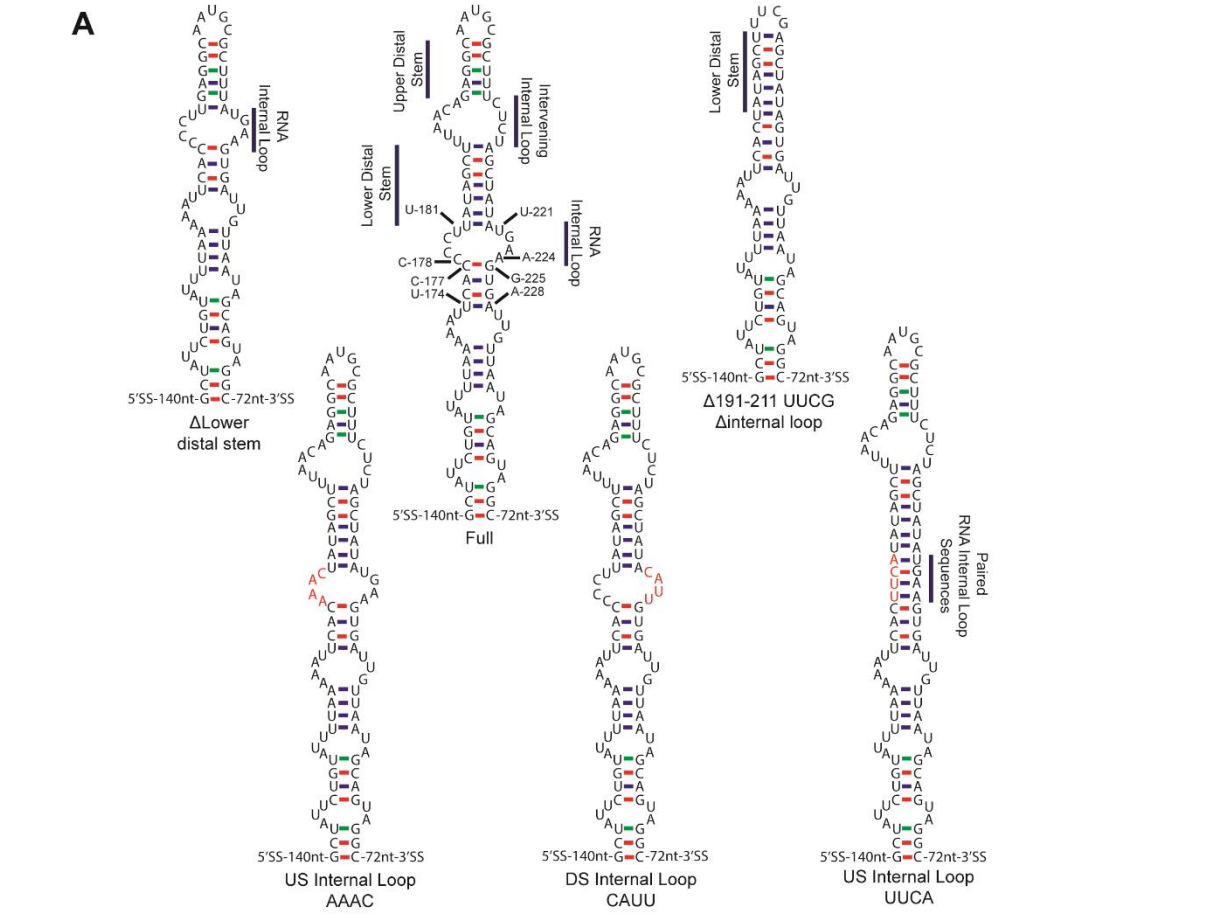

B
Scale
chrVI:
Protein-Coding Genes from Saccharomyces Genome Database
S. cerevisiae mRNAs from GenBank
Eran Segal Regulatory Module
7 yeast Multiz Alignment & Conservation
Reg. Module
Conservation
sacPar
sacMik
sackud
sacBay
sacCas
sackiu
s_cerevisiae


C
5′SS-160nt-U
A-21nt-3′SS
RPL22Bi AS 5′Δ
Δ226-297

S4 Fig


A
Hairpin Loop
Upper
Intervening Internal Loop
A-191
C-212
Lower
U-181
A-223
5′SS-168nt
91nt-3′SS
Full
B
RPL22Bi AS 5′Δ
Full
Δ153-188
Δ181-191
Δ192-211
Δ212-223
US
S
GAPDH
GFP
C
UUCG Tetraloop
U-188
C-187
5′SS-168nt
94nt-3′SS
Δ191-211
UUCG
D
RPL22Bi AS 5′Δ
Full
Δ181-191
Δ191-211 UUCG
US
S
scR1
GFP
E
5′SS-168nt
94nt-3′SS
"Flip"
5′SS-168nt
94nt-3′SS
"Scramble"
F
RPL22Bi AS 5′Δ
Full
Δ181-191
Δ191-211 UUCG
UUCG "Flip"
UUCG "Scramble"
US
S
scR1
GFP

S5 Fig

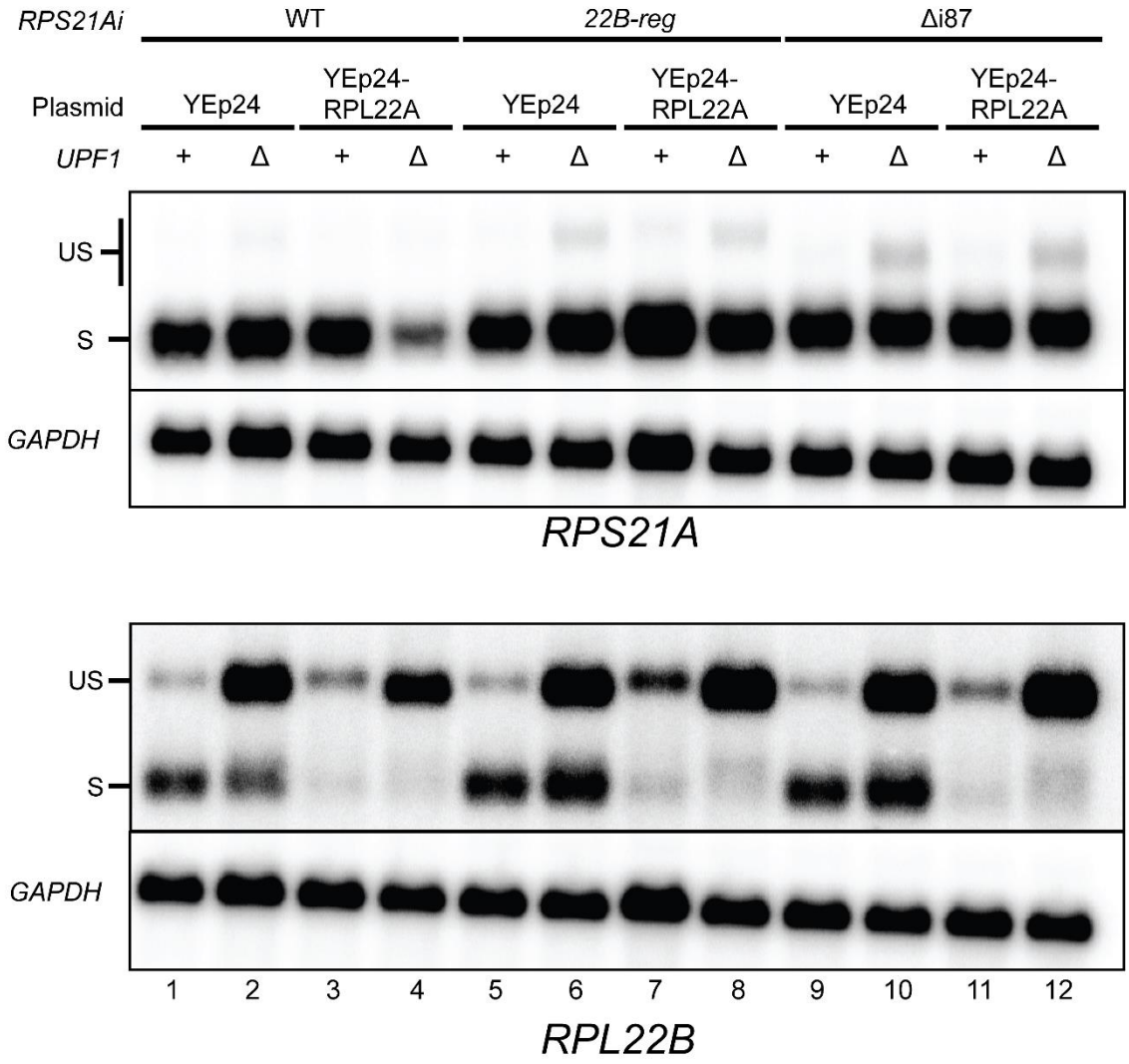

A
RNA Internal Loop
5′SS-140nt-G C-72nt-3′SS
ΔLower distal stem
Upper Distal Stem
Intervening Internal Loop
Lower Distal Stem
U-181
U-221
C-178
A-224
C-177
G-225
U-174
A-228
RNA Internal Loop
5′SS-140nt-G C-72nt-3′SS
Full
Lower Distal Stem
5′SS-140nt-G C-72nt-3′SS
Δ191-211 UUCG
Δinternal loop
5′SS-140nt-G C-72nt-3′SS
US Internal Loop
AAAC
5′SS-140nt-G C-72nt-3′SS
DS Internal Loop
CAUU
Paired RNA Internal Loop Sequences
5′SS-140nt-G C-72nt-3′SS
US Internal Loop
UUCA
B
RPL22Bi AS 5′Δ
Full
Δ181-191
Δ191-211 UUCG
Δ191-211 UUCG Δinternal loop
Δ Lower distal stem
US Internal Loop AAAC
DS Internal Loop CAUU
US
S
scR1
1
2
3
4
5
6
7
GFP
C
RPL22Bi AS 5′Δ
Full
Δ181-191
US Internal Loop AAAC
DS Internal Loop CAUU
US Internal Loop UUCA
US
S
scR1
GFP

S6 Fig

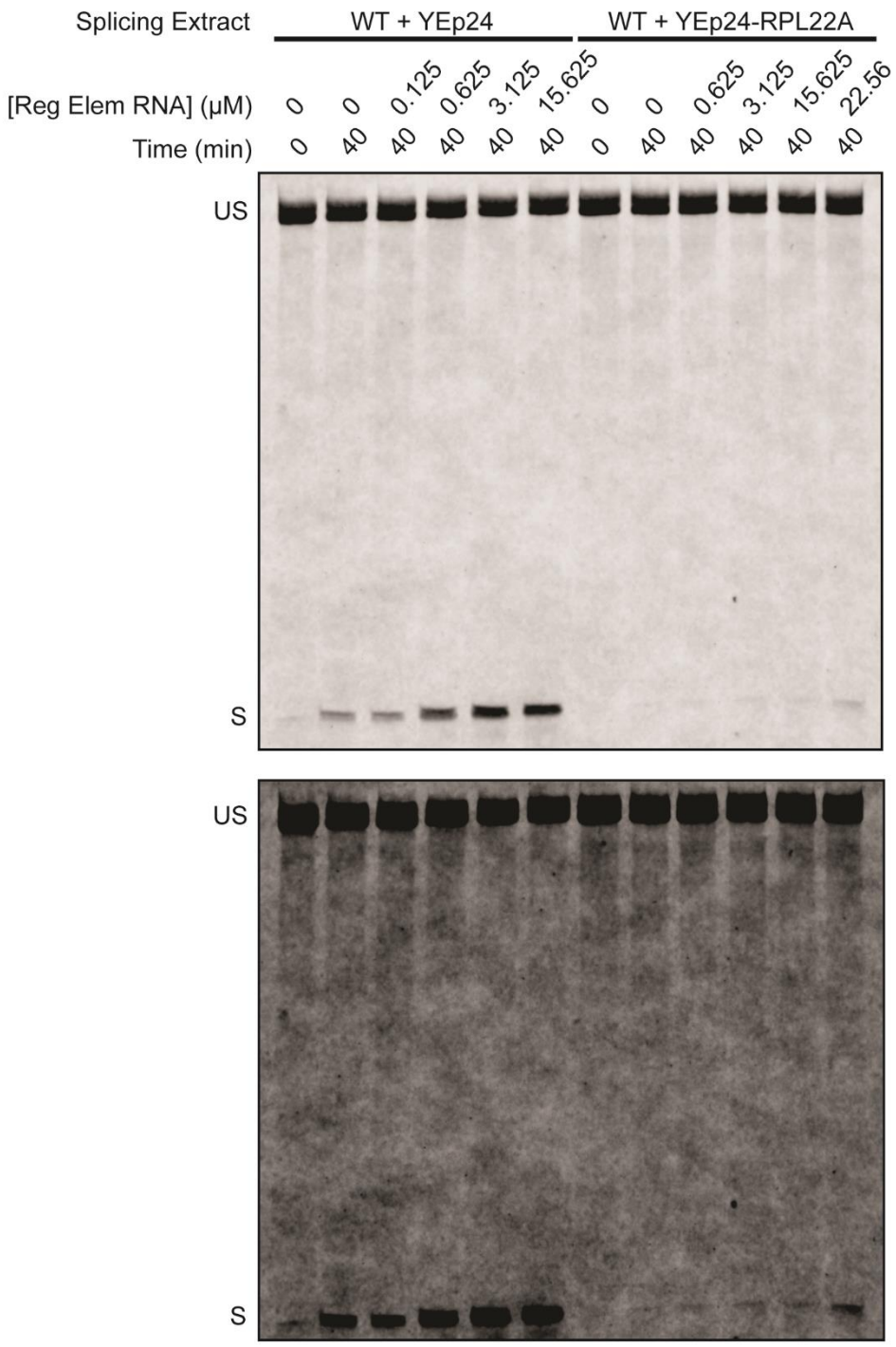
A

RPG	Intron Length	5'SS-BP Distance	Exon 1 Length
RPL22B	320	305	134
RPS21A	322	301	92

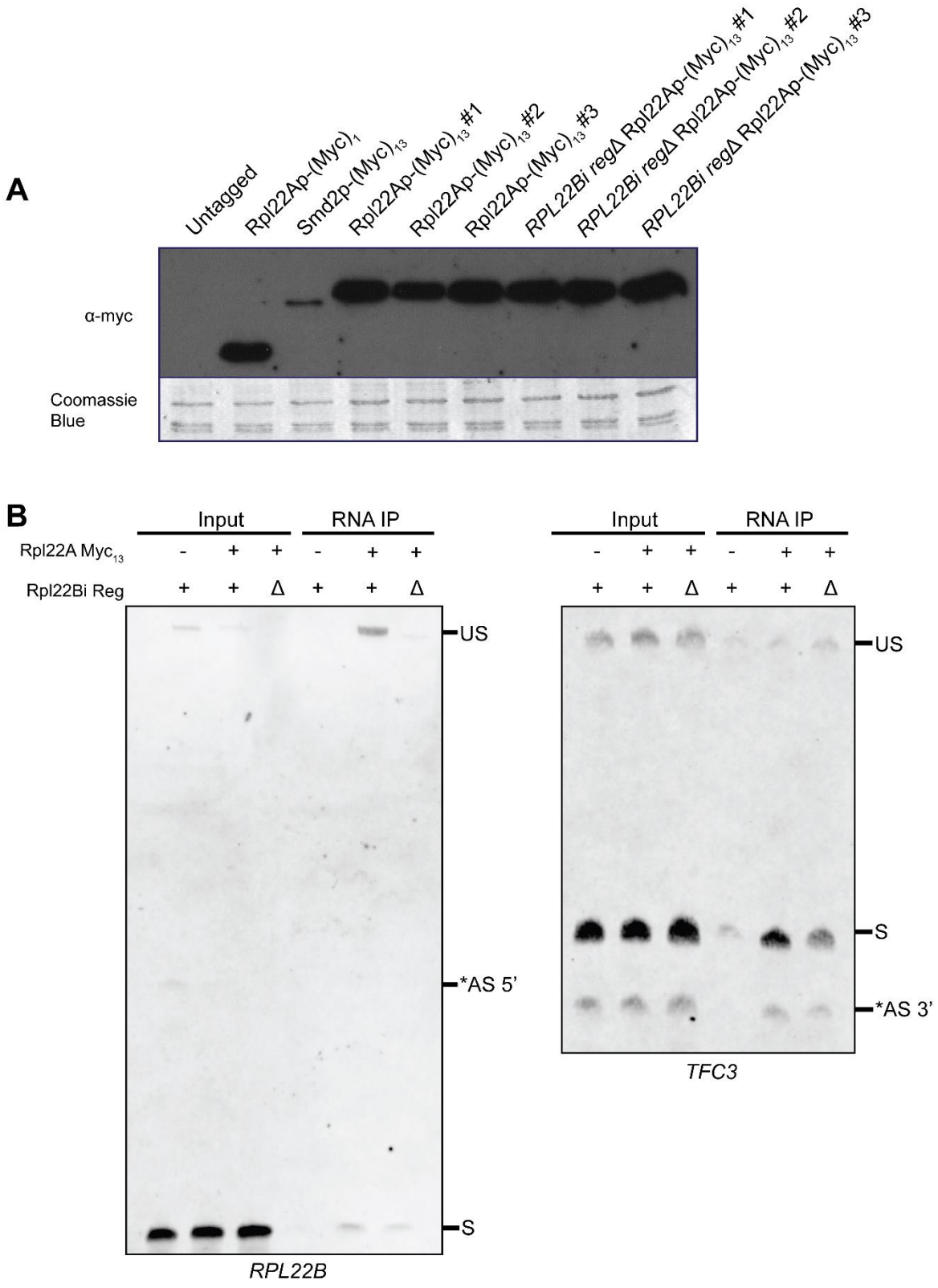
B



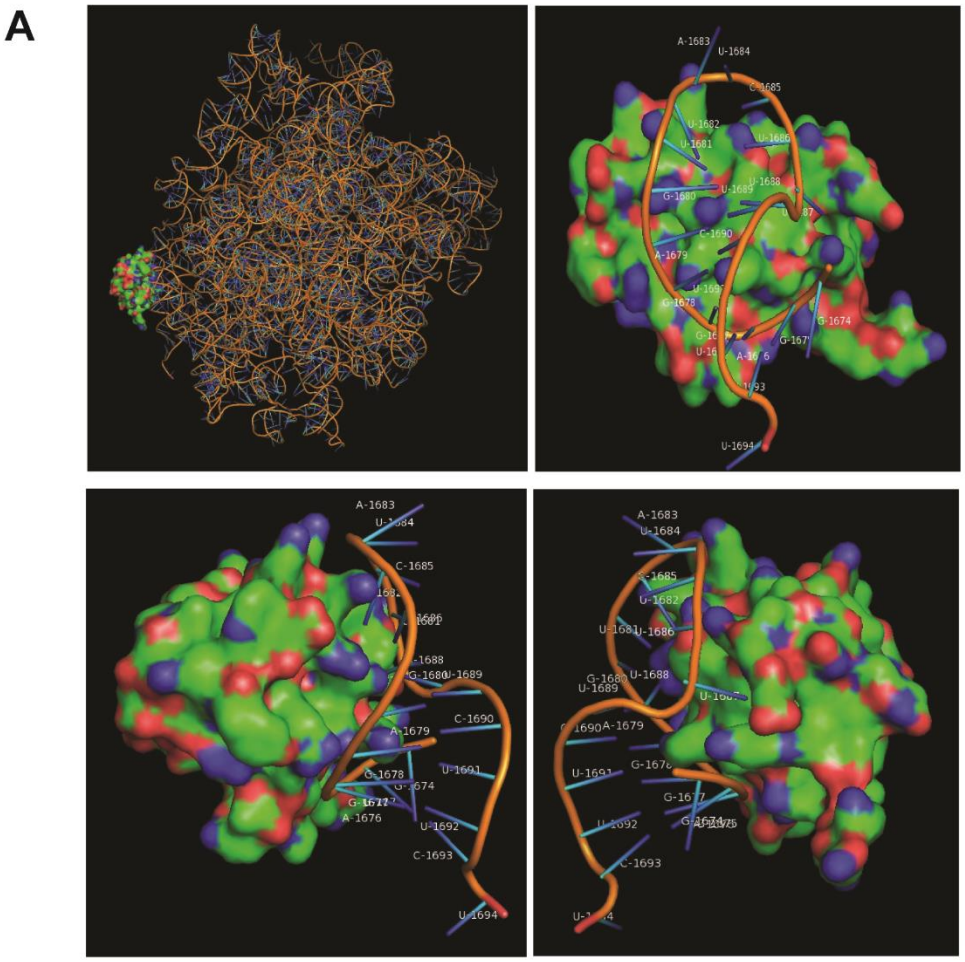
S7 Fig

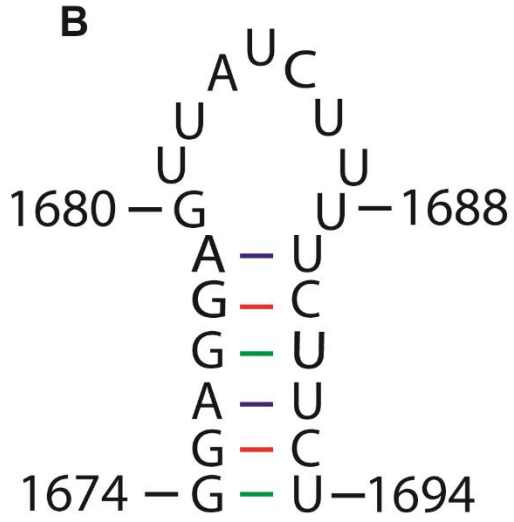

Splicing Extract
WT + YEp24
WT + YEp24-RPL22A
[Reg Elem RNA] (μM)
0 0 0.125 0.625 3.125 15.625 0 0 0.625 3.125 15.625 22.56
Time (min)
0 40 40 40 40 40 0 40 40 40 40 40
US
S
US
S

S8 Fig

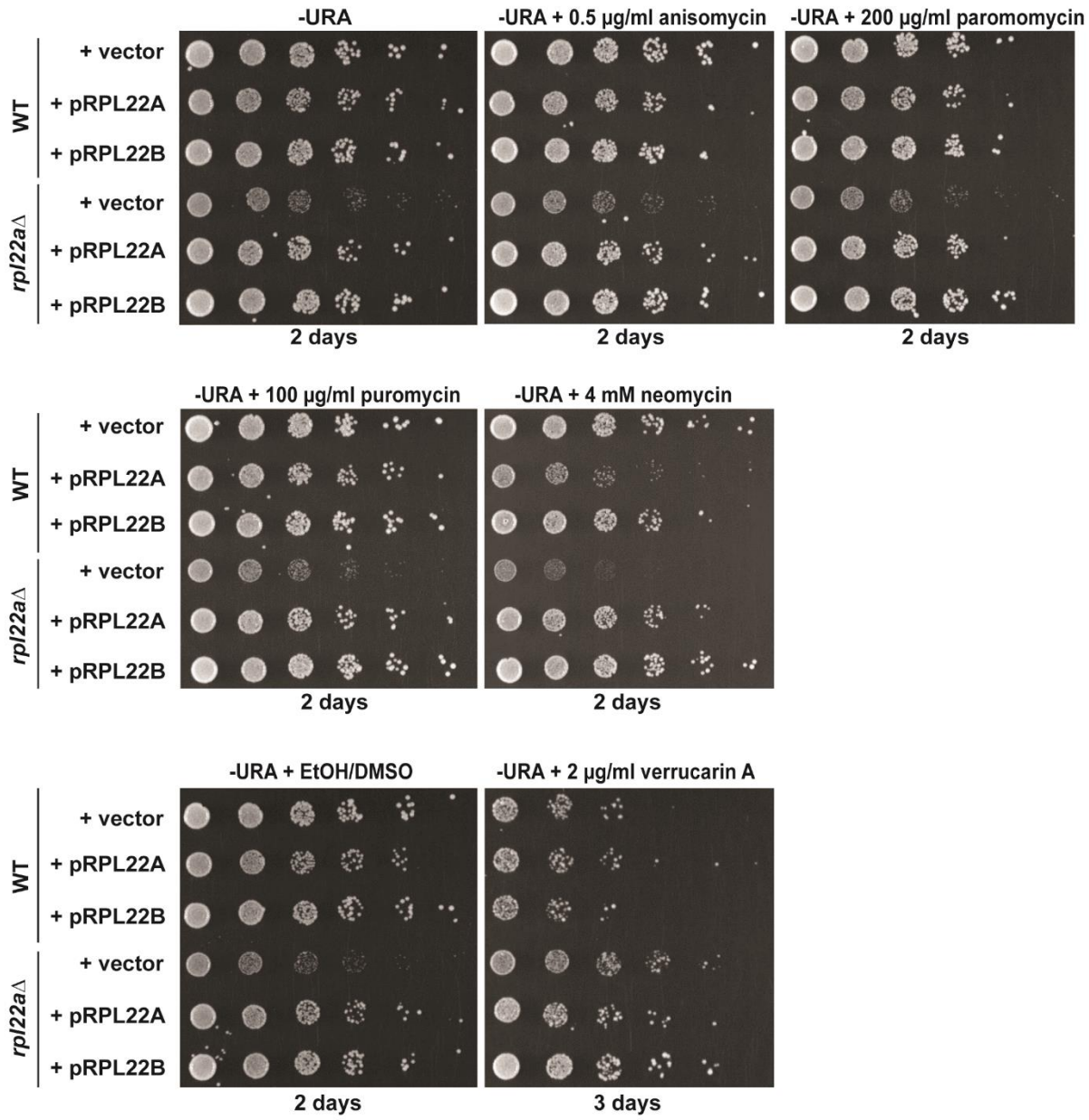

A
Untagged
Rpl22Ap-(Myc)1
Smd2p-(Myc)13
Rpl22Ap-(Myc)13 #1
Rpl22Ap-(Myc)13 #2
Rpl22Ap-(Myc)13 #3
RPL22Bi regΔ Rpl22Ap-(Myc)13 #1
RPL22Bi regΔ Rpl22Ap-(Myc)13 #2
RPL22Bi regΔ Rpl22Ap-(Myc)13 #3
α-myc
Coomassie Blue
B
Input
RNA IP
Rpl22A Myc13
Rpl22Bi Reg
US
*AS 5'
S
RPL22B
*AS 3'
TFC3

S9 Fig

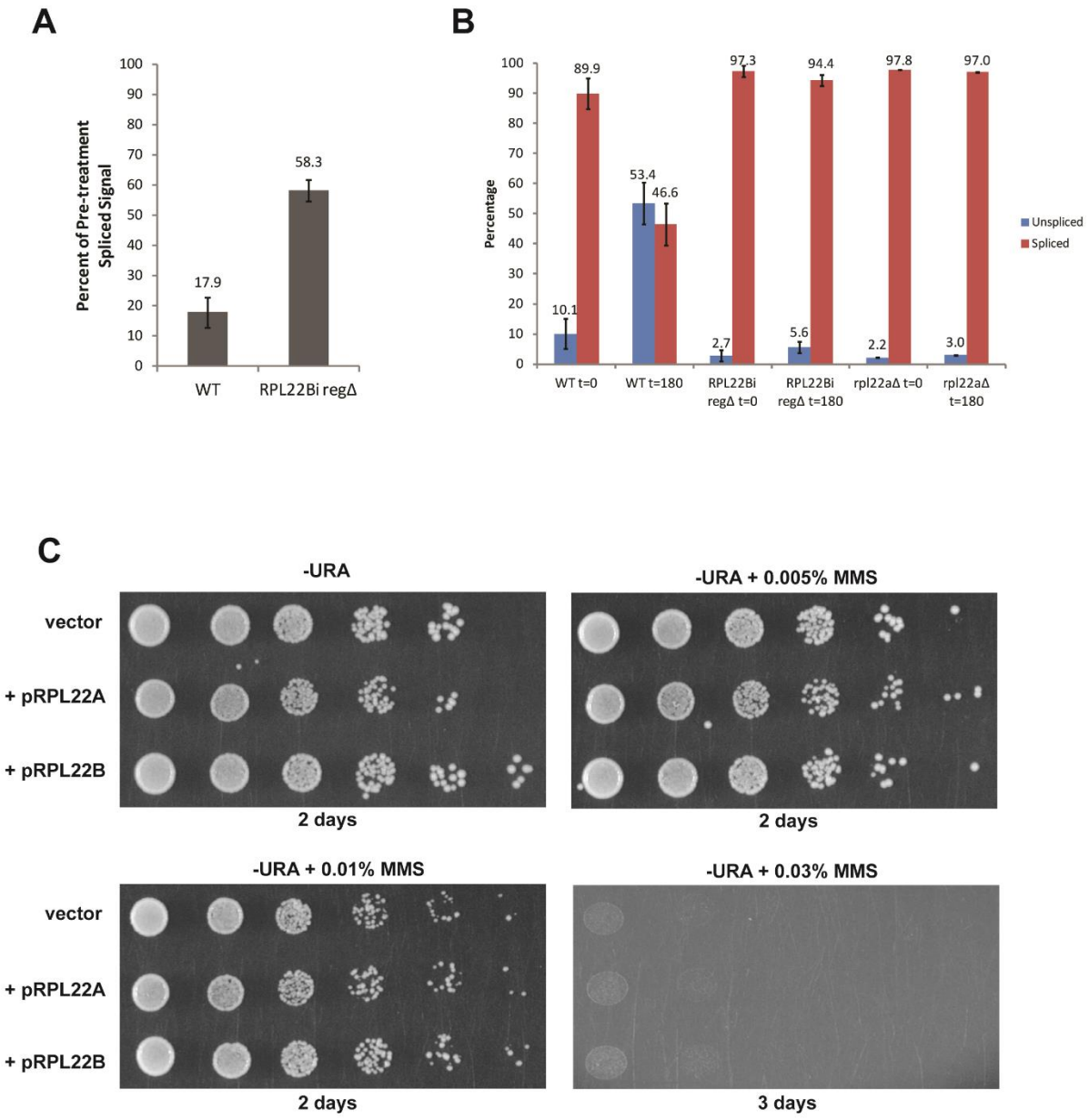

A
A-1683
U-1684
C-1685
U-1682
U-1681
U-1686
U-1688
U-1689
G-1680
C-1690
A-1679
U-1691
G-1678
G-1674
A-1676
U-1692
C-1693
U-1694


B
1680 – G
U – 1688
A – U
G – C
G – U
A – U
G – C
1674 – G – U – 1694

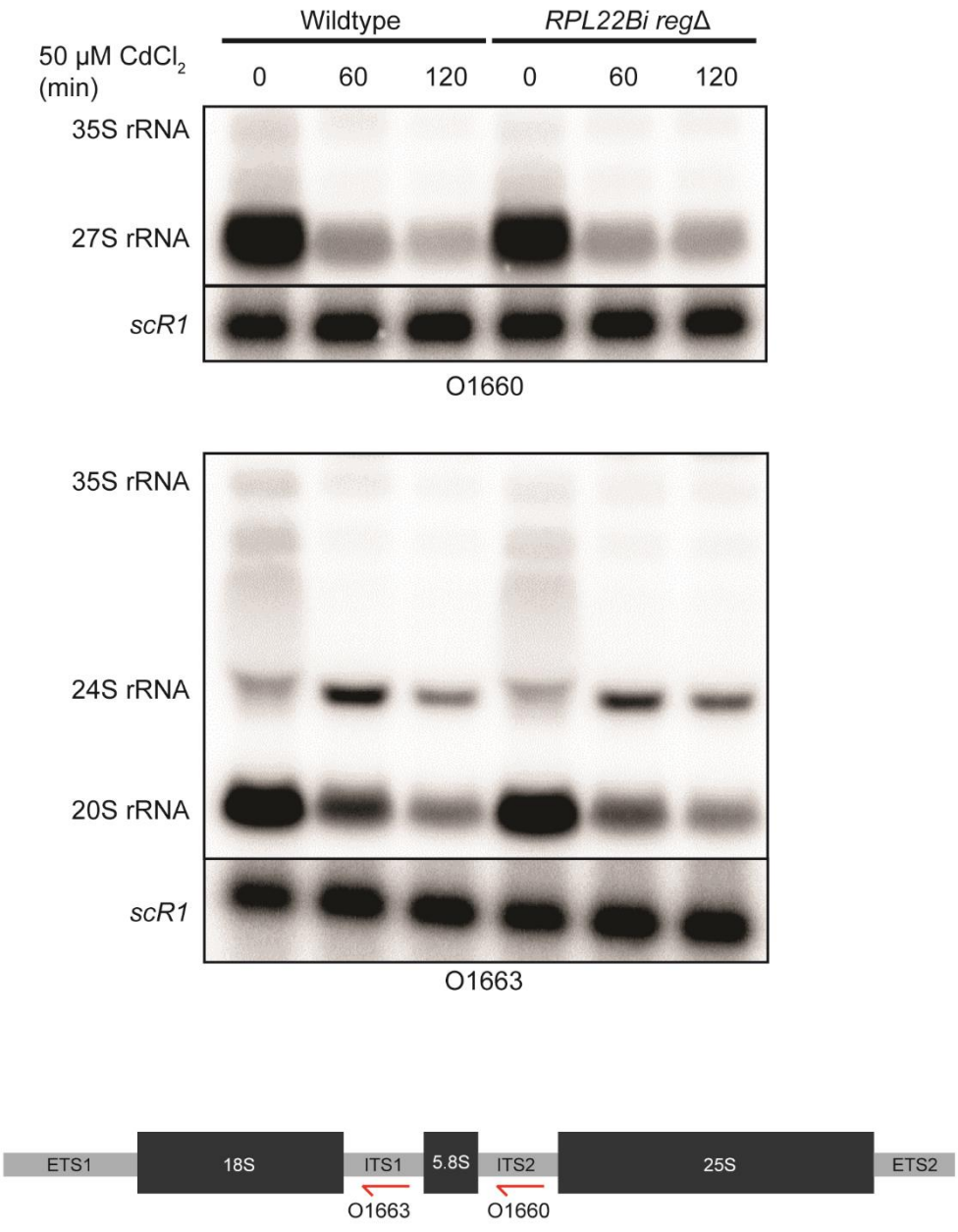
S10 Fig


-URA
-URA + 0.5 µg/ml anisomycin
-URA + 200 µg/ml paromomycin
WT
rpl22aΔ
+ vector
+ pRPL22A
+ pRPL22B
2 days
-URA + 100 µg/ml puromycin
-URA + 4 mM neomycin
-URA + EtOH/DMSO
-URA + 2 µg/ml verrucarin A
3 days

S11 Fig


A
Percent of Pre-treatment Spliced Signal
100
90
80
70
60
50
40
30
20
10
0
17.9
58.3
WT
RPL22Bi regΔ
B
Percentage
89.9
53.4
46.6
97.3
94.4
97.8
97.0
10.1
2.7
5.6
2.2
3.0
WT t=0
WT t=180
RPL22Bi regΔ t=0
RPL22Bi regΔ t=180
rpl22aΔ t=0
rpl22aΔ t=180
Unspliced
Spliced
C
-URA
-URA + 0.005% MMS
vector
+ pRPL22A
+ pRPL22B
2 days
2 days
-URA + 0.01% MMS
-URA + 0.03% MMS
vector
+ pRPL22A
+ pRPL22B
2 days
3 days

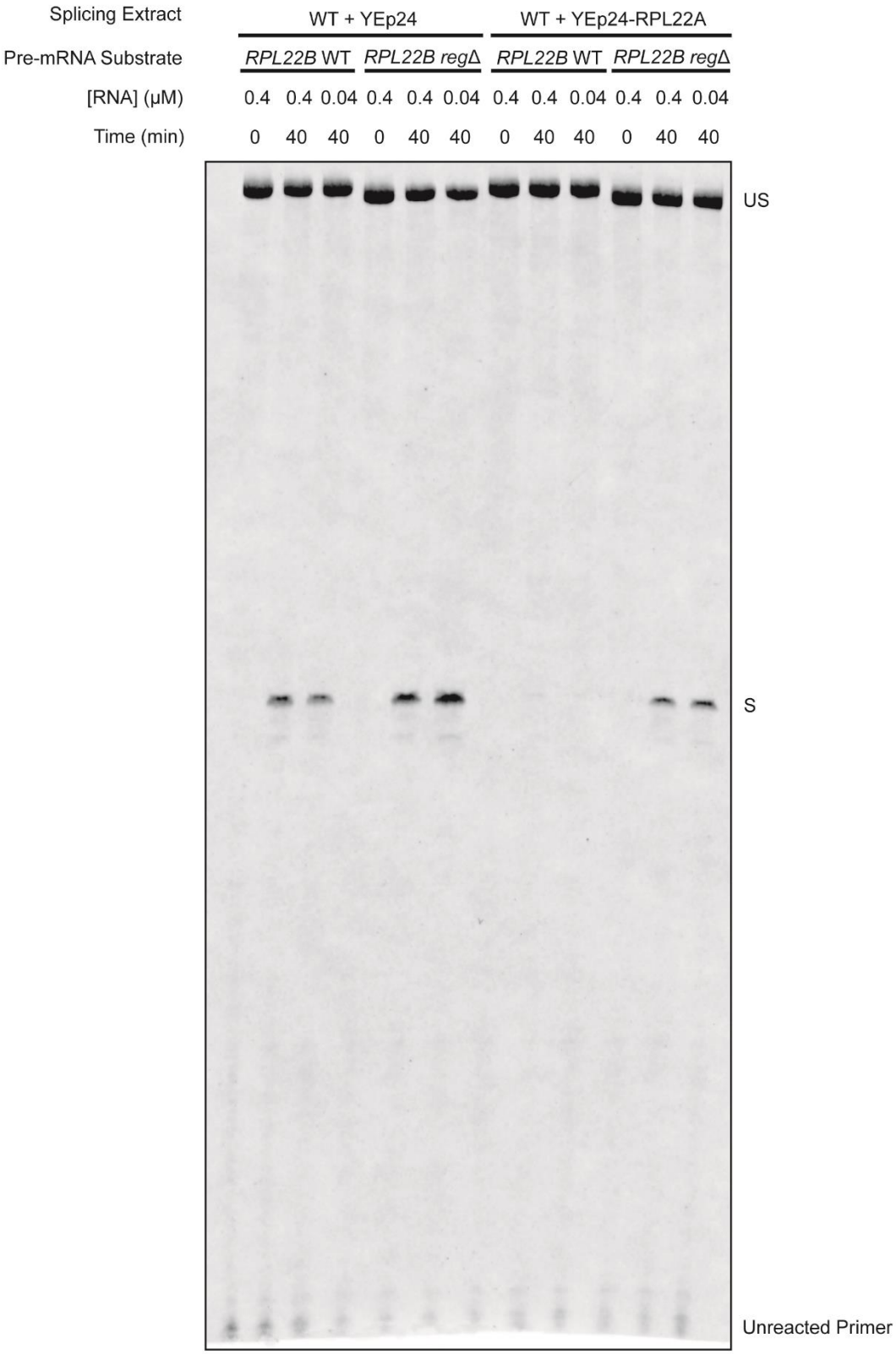
S12 Fig


Wildtype
RPL22Bi regΔ
50 μM CdCl2
(min)
0
60
120
0
60
120
35S rRNA
27S rRNA
scR1
O1660
35S rRNA
24S rRNA
20S rRNA
scR1
O1663
ETS1
18S
ITS1
5.8S
ITS2
25S
ETS2
O1663
O1660

S13 Fig


Pre-mRNA Substrate
RPL22B WT
RPL22B regΔ
[RNA] (μM)
0.24 0.24 0.024 0.0024 0.24 0.24 0.024 0.0024
Time (min)
0 40 40 40 0 40 40 40
US
S
Unreacted Primer

S14 Fig


Splicing Extract
WT + YEp24
WT + YEp24-RPL22A
Pre-mRNA Substrate
RPL22B WT
RPL22B regΔ
RPL22B WT
RPL22B regΔ
[RNA] (μM)
0.4 0.4 0.04 0.4 0.4 0.04 0.4 0.4 0.04 0.4 0.4 0.04
Time (min)
0 40 40 0 40 40 0 40 40 0 40 40
US
S
Unreacted Primer

Chapter 4

Global Suppression of Non-Canonical Alternative 3´-Splice Sites by Prp18p

Global Suppression of Non-Canonical Alternative 3´-Splice Sites by Prp18p

Jason Gabunilas, Kevin Roy, Guochang Lyu, Joyce Samson, Jonelle White, Samantha Edwards, and Guillaume Chanfreau

Abstract

The fidelity of splice site selection is essential to properly express genetic information. The Prp18p splicing factor is recruited to the spliceosome during the second step of pre-mRNA splicing and facilitates the second transesterification reaction. Although Prp18p is not essential for splicing under normal conditions, its function in 3´-splice site positioning suggests that it may impact the fidelity of 3´-splice site selection. Using RNA sequencing of Prp18 knockout strains we show that the absence of *PRP18* triggers splicing at branchpoint-proximal alternative 3´-splice sites in over 30% of all intron-containing genes. These upstream alternative 3´-splice sites exhibit a diverse array of non-consensus nucleotide sequences. A mutation in Slu7 which decreases its affinity for Prp18 and mutations at the Prp8•Prp18 interface result in the activation of the same non-consensus splice sites, emphasizing the importance of Prp18 recruitment for proper 3´-splice site selection. These results combined with recent structures of the spliceosome suggest that in the absence of Prp18, alternative, upstream 3´-splice site-like sequence can prematurely enter the spliceosome active site and compete for exon ligation before the proper 3´-splice site can be used. Taken together, these results demonstrate the importance of Prp18p in ensuring the fidelity of 3´-splice site selection and the accuracy of pre-mRNA splicing.

Introduction

The second catalytic step of pre-mRNA splicing relies on the concerted action of spliceosomal snRNPs and associated splicing factors to efficiently recognize the appropriate 3´-

splice site (3´-SS) and ligate the two exons. Following the ATP-dependent rearrangement of the spliceosome and the ejection of Yju2p and Cwc25p from the spliceosome by Prp16p after the first catalytic step (Schwer and Guthrie, 1992), a trio of second step splicing factors consisting of Slu7p, Prp18p, and Prp22p sequentially joins the spliceosome to facilitate the second catalytic step (James et al., 2002). Prp18p and Slu7p are necessary for the efficient docking of the 3´-SS into the spliceosomal active site (Ohrt et al., 2013), while Prp22p proofreads the ligated exons in an ATP-dependent manner following the second transesterification reaction (Mayas et al., 2006).

During the second step of splicing, the U5 snRNP aligns the two exons for ligation within the spliceosome by pairing the invariant nucleotides in loop 1 of the U5 snRNA with the intron-proximal exonic nucleotides (O'Keefe and Newman, 1998). Prp18p is necessary for efficient second step catalysis as pre-mRNA splicing is inhibited *in vivo* in yeast strains harboring a disrupted *PRP18* allele and in yeast splicing extracts devoid of Prp18p *in vitro* (Horowitz and Abelson, 1993). Reflecting this defect in splicing, yeast strains lacking Prp18p also display a severe growth defect in a range of different temperatures (Bacíková and Horowitz, 2002, 2005). Studies utilizing reporter constructs have shown that Prp18p plays a crucial role in stabilizing the interactions between the exonic nucleotides of mRNA splicing intermediates and the loop 1 nucleotides of the U5 snRNA (Crotti et al., 2007). In a yeast a mutant *prp18* allele in which the most highly conserved portion of the protein was removed (*prp18*Δ*CR*), the identities of these exonic nucleotides largely determine the efficiency of the second step of splicing based on how strongly they interact with the U5 loop 1 nucleotides. Specifically, sequences rich in adenosines accelerate the second step (Crotti et al., 2007). While exonic nucleotide sequences were initially only believed to impact splicing in *prp18* mutant backgrounds, further work from the Horowitz laboratory has demonstrated that they can also dictate 3´-SS selection in reporter transcripts harboring multiple competing 3´-SS in both *prp18*Δ*CR* mutants and in wild-type yeast (Crotti and Horowitz, 2009).

The role of Prp18p in stabilizing the mRNA intermediates for the second step of splicing suggests that it may also contribute to the fidelity of this step and ensure proper 3´-SS selection. We previously detected the presence of a non-canonical AUG alternative 3´-SS in the transcript *GCR1* in yeast strains deficient in Prp18p, and that this alternative splicing event generates a nonsense transcript that is subject to degradation by the nonsense mediated mRNA decay (NMD) pathway (Kawashima et al., 2014). In this study, we performed RNA-seq analysis of *prp18* null mutants and found that non-canonical alternative 3´-splice sites are utilized in a large number of intron-containing genes in the absence of Prp18p. Interestingly, we find that the activated alternative 3´-SS are more proximal to the predicted branchpoints than the annotated sites. Utilizing the endogenous *NYV1* and RAD14 transcripts to detect the activation of non-canonical alternative 3´-SS, we show that most second step splicing factors do not appear to impact 3´-SS selection in the same manner as Prp18p. However, one specific mutant allele of *PRP22* also fails to suppress non-canonical 3´-SS, suggesting that splicing fidelity is a product of both accurate splice site selection and proofreading of ligated exons. Additionally, we demonstrate that the ability of Prp18p to suppress alternative 3´-splice site selection may depend in part on its interaction with Prp8p and on its recruitment to the spliceosome by Slu7. These results highlight the importance of Prp18p in preserving the fidelity of pre-mRNA splicing.

Results

High-throughput RNA-sequencing reveals activation of splicing at non-canonical alternative 3´-SS in yeast lacking Prp18p

Our previous splicing analyses in the *prp18*Δ null mutant strain targeted a selected array of pre-mRNA transcripts, revealing the activation of a non-canonical alternative 3´-SS in *GCR1* in the absence of Prp18p that became more apparent when NMD was disabled via the genetic inactivation of the key NMD factor Upf1p (Kawashima et al., 2014). We hypothesized that the

suppression of these non-canonical sites might also occur with other pre-mRNAs at a global scale. We therefore performed a whole-transcriptome RNA-seq analysis of the wild-type, *upf1*Δ, *prp18*Δ, and *prp18*Δ*upf1*Δ yeast strains to identify splice junctions that are activated in these mutants. We obtained over 40 million uniquely mapped reads for each strain, with 3-5% of these reads mapping to splice junctions (Fig. 4.1A) Satisfyingly, we found that the number of reads mapping to unannotated alternative 3´-SS increased substantially in the *PRP18* and NMD mutants, with the *prp18*Δ*upf1*Δ double mutant showing the highest number of reads allocated to those sites (Fig. 4.1B). The double mutant also presented the greatest number of reads corresponding to non-canonical non-AG alternative 3´-SS (Fig. 4.1B). Analysis of these junctions by sequence identity showed that the absence of Prp18p activates a diverse array of non-canonical 3´-SS (Fig 4.1C). Importantly, many of these events are detected with the greatest frequency when both NMD and *PRP18* are inactivated, suggesting that a substantial number of these *PRP18* mutant-dependent alternative splicing events give rise to nonsense transcripts.

Unsurprisingly, when we considered all alternative 3´-SS activated in the *prp18*Δ*upf1*Δ mutant, we found that A and G are the most common nucleotides in the penultimate and 3´-most positions of these splice sites, reflecting the consensus sequence for annotated 3´-SS (Fig. 4.1D). However, we observed very poor conservation of the preceding pyrimidine for the alternative sites, consistent with our previous findings in NMD mutants (Kawashima et al., 2014). Interestingly, an aggregate analysis of the non-AG alternative 3´-SS revealed that the most common nucleotides at the final positions are A and C (Fig. 4.1D), signifying a potential preference of the Prp18p-devoid spliceosome for particular non-consensus sequences. However most of the non-YAG sites activated in the absence of Prp18 still maintain at least one of the two conserved residues (either the penultimate A, or the last intronic G, or both) in their sequence, showing that some of elements of the normal sequence must still be maintained. The preference for guanosine at the last position in particular may involve non-Watson-Crick interactions that this nucleotide

forms with the G_1 guanosine of the 5´-SS in order to position the 3´-SS for exon ligation (Parker and Siliciano, 1993).

We further investigated the characteristics of these non-AG alternative 3´-SS sequences, as they represent signals that differ substantially from the YAG consensus sequence (Kupfer et al., 2004; Sapra et al., 2004). We first considered the linear positioning of these alternative sites relative to the annotated 3´-SS. When weighted by normalized read counts, the annotated 3´-SS in the *upf1*Δ and *prp18*Δ*upf1*Δ mutants are positioned a median distance of 39 nt downstream of the predicted branchpoint sequences (Fig. 4.2A). However, analysis of the non-AG alternative 3´-SS in the *upf1*Δ mutant revealed that these sites are located 10 nt closer to the branchpoint than the annotated sites. Strikingly, the median distance from non-AG alternative 3´-SS to the branchpoint is further reduced by an additional 8 nt when Prp18p and Upf1p are both absent (Fig. 4.2A). Taken together, these findings suggest that Prp18p normally suppresses the use of branchpoint proximal non-canonical alternative 3´-SS that in many cases give rise to nonsense transcripts.

We sought to address the possibility that splicing at alternative 3´-SS in *prp18*Δ mutants may simply result from an increased rate of indiscriminate spliceosomal errors. Indeed, both *in vivo* and *in vitro* studies have shown that splicing of reporter transcripts is severely hindered in the absence of Prp18p (Crotti et al., 2007; Zhang and Schwer, 1997). The reduced splicing kinetics may allow more opportunities for the spliceosome to utilize incorrect splice sites. If alternative 3´-SS selection were the result of random spliceosome errors, then one might predict alternative splicing events to occur at a relatively consistent frequency among all intron-containing genes (ICGs). Indeed, an analysis of putative non-AG alternative 3´-splicing events detected by our bioinformatics analysis indicated that 58.9% of these alternative sites are found within ribosomal protein gene (RPG) transcripts, reflecting the fact that RPGs comprise 41.5% of ICGs (Fig. 4.2B) (Ares et al., 1999). However, when we considered only genes in which non-AG

alternative 3´-SS are utilized in at least 1% of all splicing events for their respective introns, we found that only 32.1% of non-AG alternative sites are found within RPGs (Fig. 4.2B). Importantly, intron-containing RPGs contribute 90% of the transcripts produced from ICGs (Ares et al., 1999). Therefore, non-RPG transcripts that are generally expressed at lower levels tend utilize non-AG alternative 3´-SS more frequently than more abundant transcripts, in agreement with our previous observations in yeast NMD mutants (Kawashima et al., 2014). In light of these findings, we propose that alternative splicing in *prp18*Δ strains is not the result of random spliceosome errors, but instead may be governed by other properties of the spliceosome or of the pre-mRNA transcript (see Discussion)

The absence of Prp18p results in the activation of upstream non-YAG 3´-splice sites in the *NYV1*, *RAD14*, *PHO85*, *MAF1* and *BET4* genes

Guided by the results of our high-throughput sequencing experiments, we validated several instances of non-canonical alternative 3´-SS activation in the *prp18*Δ mutant. The pre-mRNA transcripts for *NYV1* and *RAD14* are spliced at non-canonical 3´-SS following the inactivation of Prp18p (Fig 4.3A-B). As shown in Fig. 4.3B, the abundance of these alternatively spliced transcripts increases when NMD is disabled, in agreement with the prediction that they would trigger premature translation termination. *NYV1* displayed a particularly strong *prp18*Δ-dependent alternative splicing phenotype utilizing a highly unusual CAU alternative 3´-SS (Fig. 4.3A-B). We also validated the activation of upstream alternative 3´-SS in the transcripts for *PHO85*, *MAF1* and *BET4* as predicted by our informatics analysis (Fig. 4.4). However, while these alternative splicing events all utilized non-canonical 3´-SS sequences, the *NYV1* alternative 3´-SS was the only confirmed case that did not end in a guanosine nucleotide, illustrating a drastic departure from consensus. By contrast, all annotated 3´-SS in *S. cerevisiae* end in a guanosine (Ma and Xia, 2011). Due to the very unusual nature of its alternative 3´-SS sequence and

remarkably strong activation in the absence of Prp18p (Fig. 4.3B), we utilized *NYV1* as an endogenous model transcript to further dissect the molecular mechanisms governing non-canonical splicing in *PRP18* mutants.

Helix 5 of Prp18 but not Helix 2 or the Conserved Region is critical to suppress the use of non-canonical splice sites.

The yeast Prp18 protein consists of five alpha helices interspersed with short loop regions, including a loop connecting helices 4 and 5 that contains a highly conserved stretch of 25 amino acids (Bacíková and Horowitz, 2005; Jiang et al., 2000). Extensive mutational analysis of *PRP18* in various regions of the protein suggested two distinct roles for Prp18p – stabilizing Slu7p within the spliceosome and promoting the second step of splicing independently from its interactions with Slu7p (Bacíková and Horowitz, 2002). Several of these mutations result in growth defects (many of them temperature-sensitive) or disrupt the interaction between Prp18p and Slu7p.

We wondered whether one or more of these mutations would also impact splicing fidelity and give rise to the alternative 3´-SS events observed in the *prp18*Δ null mutant, as this would reveal which regions of the protein are necessary to ensure splicing fidelity of 3´-SS selection. To address this, we explored a panel of mutants that were similar or identical to those examined by Bacikova and Horowitz: 1) a double point mutation in helix 2 (*prp18-h2*, R151E R152E) that renders Prp18p unable to bind Slu7p, 2) a quadruple point mutation of four conserved residues in helix 5 that severely impacts growth at 30°C but retains Slu7p interaction (*prp18-h5*, D223K E224A K234A R235E), and 3) a modified deletion of the conserved loop that disrupts growth at temperatures above 34°C but binds Slu7p efficiently (*prp18*-Δ*CRt*, Δ(S187-I211)) (Bacíková and Horowitz, 2002). These mutations were introduced into a plasmid-borne copy of *PRP18* and expressed in the *prp18*Δ*upf1*Δ null mutant.

Consistent with previous observations, the *prp18-h5* mutant displayed a severe growth defect at 30°C while the growth of the other two mutants was unaffected at this temperature (Fig. 4.5A). We next analyzed the splicing patterns for *NYV1* at 30°C among these mutant strains, finding that the *prp18-h5* mutant produces the alternative 3´-SS product that is seen in the *prp18*Δ strain (Fig. 4.5B), signifying that these specific residues of helix 5 are essential to achieve full functionality of Prp18p at 30°C, including the suppression of alternative splicing. On the other hand, the alternative splice product did not appear in the *prp18-h2* mutant (Fig. 4.5B). Finally, the *prp18-*Δ*CRt* mutant also showed no signs of non-canonical alternative splicing for *NYV1*, suggesting that the conserved loop is not necessary to suppress the alternative 3´-SS. Overall, these results suggest that individual segments of Prp18p differ in their propensity to influence 3´-SS selection, with the mutation of helix 5 having the greatest impact on the growth and splicing phenotypes.

Non canonical 3´-SS usage is highly specific to Prp18 mutants

We next explored whether the non-canonical 3´-SS activated in *prp18* mutants can also be detected in other second step splicing factor mutants. The latter case would suggest a more general cooperative mechanism for maintaining splicing fidelity during the second step. We selected a panel of factors based on their roles in the second step of splicing. Prp17 stabilizes the C* complex following remodeling by Prp16 but preceding exon ligation (Fica et al., 2017). Prp22p, a DEAH-box RNA helicase which is tasked with the ejection of Slu7p and Prp18p and mRNA release following the second transesterification, also has a demonstrated role in proofreading exon ligation (Mayas et al., 2006). Cwc21p was also shown to influence 3´-SS selection under certain conditions (Gautam et al., 2015). Finally, we also analyzed mutants of the DEAH-box RNA helicase Prp43p, which operates late in the splicing process to facilitate the

disassembly of the U2, U5, and U6 snRNPs from the excised intron lariat (Fourmann et al., 2017; Tsai et al., 2005).

RT-PCR analysis of *NYV1* and *RAD14* splicing in null mutant strains for *PRP17* and *CWC21* revealed no evidence of the non-canonical alternative splicing events attributed to the absence of Prp18p for either transcript (Fig. 4.3B). Similarly, when we shifted a heat-sensitive *prp22* mutant to the non-permissive temperature, we observed a slight accumulation of the *NYV1* pre-mRNA when NMD was inactivated, similar to the *prp18*Δ null mutants, but the alternative splice product was not detected under these conditions (Fig. 4.6A).

To evaluate the temporal nature of the alternative 3´-SS selection relative to spliceosome disassembly, we analyzed mutants of the spliceosome dissociation factor Prp43p containing two different mutations at glutamine 423 within motif VI. The mutants *prp43-Q423N* and *prp43-Q423E* both demonstrate cold sensitivity *in vivo* and the *prp43-Q423E* mutant was shown to enhance splicing at non-canonical and cryptic 3´-splice sites in reporter transcripts (Mayas et al., 2010), suggesting a role for Prp43p in promoting the accuracy of exon ligation. Surprisingly, however, we found no evidence of splicing at the *NYV1* alternative 3´-SS in either of these mutants in the *upf1*Δ background at either the permissive or non-permissive temperatures (Fig. 4.6B). This result indicates that, at least for the *NYV1* transcript, the selection of the non-canonical 3´-SS likely occurs prior to the initiation of Prp43p-induced discard of the spliceosome (J. Staley, personal communication, 2017).

3´-splice site fidelity mutants in *PRP22* but not *PRP8* can activate PRP18-dependent non-canonical 3´-splice sites

Prp8p, a central component of the U5 snRNP and U5-U4/U6 tri-snRNP, is the largest spliceosomal protein, demonstrates a high degree of sequence conservation across eukaryotes, and is involved in both the first and second catalytic steps of splicing (Grainger, 2005; Horowitz,

2012). *PRP8* mutations have been described that impact the fidelity of 3´-SS selection to varying degrees (Umen and Guthrie, 1996). We selected two of these mutant alleles, *prp8-121* and *prp8-123*, that were shown to confer increased splicing for point mutations within the YAG motif of the 3´-SS of a reporter transcript. To test whether these *prp8* alleles might activate non-canonical alternative 3´-SS for endogenous transcripts, we introduced these same mutations into the chromosomal *PRP8* gene using the CRISPR-Cas system (Dicarlo et al., 2013). Neither of these mutations triggered alternative splicing at the non-canonical *NYV1* 3´-SS as observed in the *prp18*Δ mutant (Fig. 4.7A). It is possible that these *prp8* mutations may affect other transcripts but they do not seem to impact 3'-SS selection for *NYV1*.

In addition to Prp8, Prp22p was shown to proofread exon ligation (Mayas et al., 2006), as evidenced by an increase use of RAG 3´-SS in in two cold-sensitive *prp22* mutants harboring mutations within the conserved motif VI, *prp22-R805A* and *prp22-G810A* (Schwer and Meszaros, 2000). We thus analyzed NYV1 splicing in these two mutants in the *upf1*Δ NMD mutant background. Interestingly, the *prp22-G810A* mutant activated the *NYV1* alternative 3´-SS at the permissive temperature of 30°C, albeit to a lesser extent than the *prp18*Δ strain (Fig. 4.7B, left panel, lanes 2 and 5). When we analyzed the splicing of *RAD14* in these same strains, we likewise found that *prp22-G810A* displays a similar but milder phenotype to that seen in the absence of Prp18p (Fig. 4.7B, right panel, lanes 2 and 5).

To determine whether this might be exacerbated at reduced temperatures, perhaps in a manner reflective of the growth phenotypes for the *PRP22* mutants at those temperatures (Schwer and Meszaros, 2000), we also analyzed the *NYV1* splicing patterns for these mutant strains following a shift to 25 or 16°C. While the growth of the *prp22-R805A* mutant is clearly inhibited at 25°C (Fig. 4.8A), this did not appear to impact the fidelity of splicing for the *NYV1* transcript (Fig. 4.8B). Furthermore, the shift to 16°C did not activate the alternative splice site in the cold-sensitive mutants but instead resulted in a general inhibition of splicing as evidenced by

the reduction of all spliced species (Fig. 4.8B). This finding suggests that the growth defect exhibited by these mutants at reduced temperatures is likely rooted in the inability to efficiently splice pre-mRNAs in general rather than an issue with splicing fidelity. Nevertheless, the finding that similar non-canonical splicing events can be detected in both Prp18 and Prp22 fidelity mutants suggest that the two proteins cooperate in ensuring the fidelity of 3´-SS selection.

Mutations at the Prp8p•Prp18p interface disable Prp18p function in preventing the use of non-canonical 3´-splice sites

Prp8p exhibits genetic interactions with most other second-step splicing factors including Prp18p (Horowitz, 2012). Furthermore, certain mutants of *PRP18* have been shown to suppress mutations in *PRP8* that confer temperature sensitivity (Aronova et al., 2007). In light of this information, wondered whether the function of Prp18p to suppress non-canonical splice sites depends on a physical interaction with Prp8p. Recent structural studies of the yeast catalytically active spliceosome have shown that Prp18p and Prp8p interact through the latter's RNase H-like domain during the second step of splicing (Fig. 4.9A) (Yan et al., 2016). We first introduced two mutations into the endogenous *PRP8* gene at P1984 and L1988 that were shown to contact residues N190 and Q181 of Prp18p, respectively (Fig. 4.9A) (Yan et al., 2016). This *prp8-P1984A-L1988A* mutant showed a very mild accumulation of unspliced *NYV1* transcript as well as a minor activation of the alt. 3´-SS that was observable even with functional NMD (Fig. 4.9B, lanes 5 and 6).

To further test for a role of Prp8 in recruiting Prp18 to ensure proper use of canonical splice sites, we mutated a plasmid-borne *PRP18* allele at Q181 and N190 and analyzed the effect of this mutation on the growth of Prp18p-deficient yeast as well as the splicing pattern of *NYV1*. As shown in Fig. 4.9C, the *prp18-Q181A-N190A* allele fully rescued the severe growth defect that is normally associated with the *prp18*Δ*upf1*Δ double mutant, suggesting that this allele largely

restores spliceosome functionality. Surprisingly, however, this allele failed to suppress splicing at the Prp18p-specific non-canonical alternative splice site for *NYV1* (Fig. 4.9D). These results indicate that spliceosome functionality sufficient for normal cell growth is attainable even when Prp18p is unable to suppress non-canonical alternative splicing for all transcripts, thus distinguishing the alternative splicing and growth phenotypes of *PRP18* mutants. Importantly, this result demonstrates that mutating the Prp18•Prp8 interface can trigger the activation of non-canonical 3´-SS, emphasizing the importance of the recruitment of Prp18 by Prp8 to ensure its function.

The relatively minor alternative splicing phenotype observed with the *PRP8* mutations compared to the reciprocal mutations in *PRP18* is potentially due to the nature of the hydrogen bonds between the pertinent residues – this interaction involves the amino acid side chains of Prp18p Q181 and N190 and the main chain backbone atoms of Prp8p P1984 and L1988. Therefore, the mutation of the involved Prp8p residues to alanine would be expected to have a lesser effect on the interaction, as the side chains of P1984 and L1988 are not involved (Fig. 4.9A).

SLU7 **overexpression rescues** ***prp18*****Δ growth defects, but does not suppress non-canonical alternative splicing**

In addition to Prp8, Slu7 is also involved in the recruitment of Prp18p to the spliceosome during the second step of splicing: Slu7p interacts physically and genetically with Prp18p and was shown to mediate 3´-SS choice (Frank and Guthrie, 1992). Due to their cooperative involvement with facilitating the second catalytic step of splicing, we questioned whether an overabundance of Slu7p could rescue the defects in growth and 3'-SS selection that are characteristic of the *prp18*Δ mutant. To test this, we introduced either the *PRP18* or *SLU7* genes into the multicopy YEp24 vector and transformed these constructs into the null *prp18*Δ*upf1*Δ mutant. The

overexpression of Prp18p and Slu7p conferred complete and partial rescue, respectively, of the characteristic growth defect observed in the *prp18*Δ*upf1*Δ strain (Fig. 4.10A). The partial growth rescue observed with Slu7p overexpression in these conditions is consistent with previous observations (Bacíková and Horowitz, 2002) and could potentially be attributed to a partial restoration of global splicing efficiency. Indeed, previous studies have shown that while Prp18p-depleted yeast splicing extracts are normally severely defective for splicing *in vitro*, the addition of excess Slu7p is able to restore splicing activity (Zhang and Schwer, 1997). Surprisingly, the overexpression of Slu7p in the *prp18*Δ*upf1*Δ had no observable effect on either the utilization of the alternative 3´-SS or on overall splicing efficiency for *NYV1 in vivo* (Fig. 4.10B). This result suggests that the absence or presence of wild-type Slu7p neither activates nor suppresses, respectively, alternative 3´-SS in the same manner as Prp18p, at least for *NYV1*.

We also analyzed the splicing of other substrates in these conditions, including the ribosomal protein gene *RPL22B* and the small non-coding intron-containing RNA *snR17*, the latter of which is not subject to translation and would therefore be insensitive to NMD (Maquat, 2004). The splicing patterns for these transcripts were similarly unaffected by the overexpression of Slu7p (Fig. 4.11). We note that although Slu7p overexpression did not appear to impact the splicing of the few transcripts analyzed here, it remains possible that the partial rescue of the *prp18*Δ-mediated growth defect is dependent upon the improved splicing for a subset of pre-mRNAs that are limiting for growth.

A *SLU7* mutant partially phenocopies alternative splicing observed in *prp18*Δ

We next asked whether Prp18p requires Slu7p in order to suppress the activation of alternative 3'-SS. We first implemented the anchor away technique to deplete either Prp18p or Slu7p from the nucleus - this approach removes the target protein from the nucleus by physically coupling it to the nuclear export of the highly abundant ribosomal protein Rpl13Ap upon the

addition of rapamycin (Haruki et al., 2008). As expected, the nuclear depletion of either Prp18p or Slu7p resulted in a clear splicing defect as evidenced by the accumulation of the pre-mRNA (Fig. 4.12A). Additionally, the alternative 3´-SS band for *NYV1* was barely detectable following two hours of depleting Prp18p from the nucleus and was undetectable in the Slu7p anchor away within the context of active NMD (Fig 4.12A). However, in the absence of functional NMD, the *NYV1* alternatively spliced band intensified within 60 minutes of rapamycin addition in the Prp18p anchor away strain (Fig. 4.12B, lanes 9 and 10), indicating higher utilization of the alternative splice site upon depletion of Prp18p from the nucleus. By contrast, anchor away of Slu7p did not activate the alternative 3´-SS for *NYV1* within the same time frame (Fig. 4.12B, lanes 11 and 12).

In performing the aforementioned nuclear depletion experiments, we recognized the possibility that the anchor away targeting Slu7p might also titrate other splicing factors out of the nucleus, which might in turn strongly inhibit the overall splicing mechanism. In light of this, we analyzed the splicing phenotypes for *NYV1* in a panel of temperature-sensitive *SLU7* mutant strains consisting of *slu7-11* (H75R, R243G, D267G), *slu7-14* (R173G, W210R, L297Q), and *slu7-EIE* (E215A-I216A-E217A), in which cell growth is completely inhibited at the non-permissive temperature (Fig. 4.13A) (Aronova et al., 2007). Interestingly, the *slu7-EIE* allele was the only *slu7* mutant which exhibited usage of the non-canonical alternative splicing product, which appeared to be stronger at 30°C than at the non-permissive temperature of 37°C (Fig. 4.13B, lanes 4 and 9). This result is consistent with previous experiments showing that the *slu7-EIE* mutation abolishes the interaction of Slu7p with Prp18p (James et al., 2002). Therefore, it is possible that the protein product of *slu7-EIE* is unable to effectively recruit of Prp18p to the spliceosome, and thus partly phenocopies the absence of Prp18p. However, this was also surprising given that the *prp18-h2* mutation (see above), which likewise disrupts the Prp18p-Slu7p interaction, showed an activation of the alternative 3´-SS (Fig. 4.5B). This may suggest that the disruptive mutations on Slu7p and Prp18p affect the physical interactions between these two

proteins to different extents. We conclude that the recruitment of Prp18p by Slu7 and Prp8 is critical to maintaining the fidelity of 3´-SS selection.

Discussion

In this work, we have analyzed the transcriptome of NMD-compromised yeast strains lacking the second step splicing factor Prp18p and found that the absence of this protein results in the activation of a highly diverse array of alternative 3´-splice sites, many of them deviating significantly from the consensus sequence (Fig. 4.1C). In analyzing various *PRP18* mutants and rescue strains, we found that alternative 3´-SS selection can be genetically separated from the growth defect exhibited by strains lacking Prp18p. Notably, while certain second step splicing factor mutants analyzed in this study did not reflect the splicing fidelity defect displayed by *prp18*Δ, we identified specific mutations in Slu7p and Prp22p that activate non-canonical splicing in the endogenous *NYV1* reporter transcript, albeit to a lesser extent than the *prp18*Δ null mutant.

In the case of Slu7p, we tentatively attribute the activation of the alternative 3´-SS for *NYV1* in the *slu7-EIE* mutant (Fig. 4.13B) to the inability of this mutant to recruit Prp18p to the spliceosome, consistent with the observed order of assembly of Prp18p and Slu7p onto the spliceosome (James et al., 2002). However, the *prp18-h2* mutant, which is unable to bind Slu7p and therefore presumably also cannot be recruited to the spliceosome, does not show the same 3´-SS phenotype (Fig. 4.5B). It is possible that the *prp18-h2* mutant protein can still bind the spliceosome independently of Slu7p, perhaps with reduced affinity. Moreover, we have shown that although the overexpression of Slu7p partially rescues to the growth defect of the *prp18*Δ mutant, it fails to suppress non-canonical alternative splicing in that mutant, at least for the *NYV1* transcript (Fig. 4.10). This indicates that in addition to a general cooperative role in facilitating the second step of splicing, Prp18p and Slu7p may also exhibit unique functions in promoting splicing fidelity for different sets of RNA transcripts, some of which may be limiting for growth.

A role for Prp22p in proofreading exon ligation was previously reported (Mayas et al., 2006). Interestingly, a recent study implementing reporter splicing substrates harboring competing 3´-SS demonstrated that wild-type Prp22p enables the spliceosome to select an upstream alternative 3´-SS for exon ligation and that the ATPase activity of Prp22p is required to enable the spliceosome to recognize the downstream 3´-SS (Semlow et al., 2016). Our results further show that that upstream alternative 3´-SS of endogenous transcripts can be utilized in strains that express the cold-sensitive *prp22-G810A* mutant, even at permissive temperatures and in the presence of wild-type Prp18p (Fig. 4.7 and 4.8). Collectively, these findings support the viewpoint that, at least for a subset of transcripts, alternative splicing occurs constitutively at a basal level and that the suppression of alternative splicing or of alternatively spliced transcripts is the joint responsibility of multiple factors, including Prp18p and Prp22p.

Our transcriptome-wide analysis of alternative splicing events in *PRP18* mutant yeast revealed two intriguing properties that appear to apply to the vast majority of the activated alternative 3´-SS. First, these alternative sites are located proximal to their respective predicted branchpoint sequences relative to the annotated 3´-SS (Fig. 4.2A). While the molecular basis for this has yet to be determined, our findings are consistent with previous *in vitro* experiments showing that splicing extracts depleted of Prp18p are defective at splicing transcripts whose 3´-SS are positioned >12 nt downstream of the branchpoint, and that splicing efficiency can be restored by shortening the branchpoint-3´-SS interval to 7 nt (Zhang and Schwer, 1997). Notably, similar observations have also been recorded for mutants in the step 2 splicing factors *SLU7* and *PRP22*, both of are dispensable for splicing when the 3´-SS is sufficiently close to the branchpoint (Frank and Guthrie, 1992; Schwer and Gross, 1998). Thus, in addition to facilitating the docking of the 3´-splice site into the spliceosomal active site (Ohrt et al., 2013), it is possible that Prp18p also modulates the flexibility of the spliceosome with respect to splice site choice, perhaps by

guiding the spliceosome to the correct site, or alternatively by passively preventing the utilization of incorrect sites.

A second notable observation from our analysis is that alternative splice sites do not appear equally across all ICGs. Illustrating this point, RPGs are over-represented among all ICGs, but are under-represented among ICGs that experience the highest frequencies of alternative splicing (Fig. 4.2B). This suggests that alternative splicing in the *prp18*Δ mutants is likely not the result of aberrant spliceosome errors, but rather is governed by specific characteristics inherent to either the mutant spliceosome or the RNA transcripts themselves. Because splicing signals presumably evolved within the context of fully functional spliceosomes, it is unlikely that the alternative splice sites identified in this study serve any regulatory purpose, as they are only appreciably activated in the absence of Prp18p. One possibility is that the pseudo-exonic nucleotide sequences immediately adjacent to the activated alternative 3´-SS may contribute toward determining the utilization of their respective sites, as this has been demonstrated in reporter transcripts even for wild-type spliceosomes (Crotti and Horowitz, 2009). Surprisingly, our global analysis of 3´-SS-proximal exonic sequences did not reveal any obvious trends in this regard (Fig. 4.1D), although there does appear to be a minor enrichment for adenine nucleotides at the +1 and +2 positions. Additionally, it is possible that the impact of these exonic nucleotides on potential alternative 3´-SS selection may vary with distance from the branchpoint. Additional analyses in which the 3´-SS sequences are binned by branchpoint distance may reveal how the interplay of these two properties influences splice site selection.

Validation of the AC alternative 3´-SS detected by our informatics analysis will be a significant objective, especially because this sequence typically manifests in AT-AC introns in higher eukaryotes that possess the minor spliceosome machinery necessary to process them (Wu and Krainer, 1999). Notably, one case of an endogenous AT-AC intron embedded within the annotated intron of *RPL30* was recently reported in yeast (Gould et al., 2016). However, our

analysis suggests that in contrast to these AT-AC introns, the introns harboring AC alternative 3´-SS that are activated in the absence of Prp18p utilize the annotated 5'-splice sites. Further studies will be required to ascertain the accuracy of this prediction.

Finally, it is important to recognize that our study was restricted to the analysis of splicing of annotated introns. A substantial number of novel introns been recently identified (Schreiber et al., 2015; Volanakis et al., 2013), and it is likely that annotated yeast intronome will continue to expand as additional introns are discovered. We realize the possibility that Prp18p may also play a role in governing the splicing of these newly discovered introns as well, either in the general sense by promoting the second step of splicing, or more specifically by suppressing alternative splicing. Continuing studies should provide additional clues as to the extent to which Prp18p and other splicing factors influence the processing of these unannotated introns.

Materials and Methods

Yeast strain construction, Northern blot analysis, RT-PCR analysis, molecular cloning, restriction enzyme-based plasmid construction and site-directed mutagenesis, and spot dilution assays were performed as previously described (Gabunilas and Chanfreau, 2016; Roy and Chanfreau, 2014; Roy et al., 2016). Construction of the PRP18/pUG35 and PRP18/Ep24 base plasmids was accomplished using Gibson assembly (Gibson et al., 2009) (New England Biolabs #E2611). Unless otherwise indicated in the figure legends, all strains were derived from the BY4741 background. The *PRP8* genomic mutations were introduced using the CRISPR-Cas system as previously described (Gillespie et al., 2017). The *SLU7* and *PRP22* mutant strains and plasmids were generously provided by Beate Schwer (Cornell U.) and Jonathan Staley (U. Chicago), respectively.

Yeast Cell Culturing

Yeast strains without plasmids were grown in either YPD (1% w/v yeast extract, 2% w/v peptone, and 2% w/v dextrose) or synthetic complete media (SDC: 0.67% w/v yeast nitrogen base, 2% w/v dextrose, and 0.2% w/v powdered amino acid mixture) to exponential phase growth as previously described (Gabunilas and Chanfreau, 2016). For strains transformed with pUG35, YEp24 or their derivatives, -URA synthetic dropout media was used. For strains transformed with pRS315 or its derivatives, -LEU synthetic dropout media was used. All other growth conditions are as indicated in the figure captions. For steady-state analysis experiments, cell cultures were grown to exponential phase until an OD600 of ~0.4 – 0.6 was attained and then harvested and flash-frozen in liquid nitrogen as previously described (Gabunilas and Chanfreau, 2016; Roy and Chanfreau, 2014). For heat-sensitive mutant strains, cell cultures were first grown to exponential phase at 25°C, a portion harvested, and the remaining culture shifted to 37° and incubated at that temperature for 1 hour. For cold-sensitive mutant strains, cell cultures were first grown to exponential phase at 30°C, a portion harvested, and the remaining culture split and shifted to 16°C or 25°C and incubated at those temperatures for 2 hours. Anchor away experiments were performed as previously described (Roy et al., 2016).

Read mapping and splice junction identification

Samples from wild-type cells, and cells lacking either *UPF1*, *PRP18*, or both were purified by standard phenol-chloroform extraction. Fragmented libraries were prepared by the Illumina TruSeq stranded mRNA library prep kit. Sample quantification, sequencing and demultiplexing were performed by Macrogen/Axeq (South Korea). Reads were trimmed at the 3´ends for low quality base calls and TruSeq adapters using bbduk (Bushnell, 2014) (with settings ktrim=r k=23 mink=11 hdist=1 tpe tbo), and mapped to the R64-2-1 version of the *S. cerevisiae* genome (downloaded from the Saccharomyces Genome Database (SGD)) using STAR(Dobin et al., 2013) (with splice junction settings --outSJfilterDistToOtherSJmin 0 0 0 0 --

outSJfilterOverhangMin 6 6 6 6 --outSJfilterCountUniqueMin 1 1 1 1 --outSJfilterCountTotalMin 1 1 1 1 --chimSegmentMin 20). The outSJfilterOverhangMin flag requires that splice junctions be supported with at least 6 base pairs for each split portion of the read. The splice junction annotations from SGD were converted from the GFF to GTF format, and STAR was run in the 2-pass mode, where in the first pass reads were aligned using the annotated splice junctions. All of the junctions obtained from each sample were combined and then STAR was run again using these junctions as the "annotated" junctions.

Alternative splice site analysis and branchpoint identification

The splice junctions reported by STAR were first divided into 4 categories: junctions identical to the annotated junctions from SGD, unannotated junctions not overlapping known introns, junctions in genes with known introns and utilizing an alternative 5′ splice site (5′SS), and junctions in genes with known introns and utilizing an alternative 3′ splice site (3′SS). To avoid inclusion of low confidence junctions, a minimum intron size of 15 bp and a maximum intron size of 2000 bp were used. Each junction was first analyzed for the sequence immediately downstream of the 5′ and 3′ splice sites, as well as immediately upstream. If the sequence was identical, the junction was assigned as ambiguous, and the adjusted junction with the 5′SS signal most closely matching the consensus GTATGT was used. Mismatches in the TATGT were penalized with a score of 1 each and a mismatch in the conserved G was penalized with a score of 2. The 3 nucleotides immediately upstream of the adjusted 3′ splice site were taken as the 3′SS signal, and categorized as NAG, non-NAG, or non-NNG.

For each 3′splice site, the intronic sequence was computationally searched for the sequence most closely matching the branchpoint consensus TACTAAC, with each mismatch penalized by a score of 1, and an additional score of 2 given if the conserved TAAC was not present. The sequence with the lowest mismatch score was assigned as the most likely

branchpoint, and ties were broken by taking the branchpoint with the shortest distance to the 3´SS. The distribution of distances between the 3´SS and the branchpoint for each category NAG, non-NAG, or non-NNG was plotted using Python 2.7 and the pyplot module from matplotlib. All bioinformatics figures were prepared using Adobe Illustrator CS5. Sequence logos were generated using WebLogo 3.5.0 (Crooks et al., 2004).

Figure Legends

Figure 4.1. RNA-seq reveals the activation of alternative 3´-splice sites in the absence of Prp18p. A. Table of read properties from RNA-seq run. **B.** Histogram displaying the number of reads mapping to splice junctions that utilize any alternative 3´-splice sites (upper panel) and non-AG alternative 3´-splice sites (lower panel). Splice junction reads were normalized to the total number of reads within each strain and to total reads in WT. **C.** Stacked bar graph showing the frequency of non-AG 3´-splice site junctions by individual sequence. **D.** Sequence logo analyses of all alternative 3´-SS and non-AG alternative 3´-SS detected by RNA-seq as well as the adjacent nucleotide sequences, weighted by the number of junctions reads.

Figure 4.2. Alternative 3´-SS activated in *PRP18* mutants are branch point proximal and are activated independently of transcript expression levels. A. Histogram displaying the probability density for splicing at non-AG splice sites as a function of the distance from the branch point in the *upf1*Δ and *upf1*Δp*rp18*Δ strains. The superimposed dashed lines indicate the median distances of annotated and non-AG 3´-splice sites from the branchpoint, weighted by the normalized read counts. **B.** Comparison of the non-AG alternative 3´-splicing events identified within intron-containing RPGs and non-RPGs.

Figure 4.3. Validation of non-canonical, branch point proximal alternative 3´-splice sites for intron-containing transcripts in strains lacking Prp18p. A. Schematic diagrams of the *NYV1* and *RAD14* transcripts illustrating the splicing events that occur for each transcript. Both transcripts are spliced at the annotated (S) and alternative (*AS 3´) 3´-splice sites. *RAD14* is also spliced at two additional alternative 3´-splice sites, labeled as x and y. The *AS-3´ product for *NYV1* was confirmed by RT-PCR, cloning, and sequencing. Alternative splicing products for *RAD14* could not be sequenced due to the close spacing of the bands (see B below), but were deduced by electrophoresis gel migration pattern and locations of the –AG dinucleotides presumed to be used as splicing sequences. Numeric values indicate the number of nucleotides in the coding sequences of the exons and do not include untranslated regions. Diagrams are not to scale. **B.** RT-PCR analysis of *NYV1* and *RAD14* splicing in wild-type and spliceosome mutant strains. Band labels indicate unspliced (US) and spliced products as depicted in the diagrams in A. The asterisk in the *NYV1* panel denotes an unidentified NMD-sensitive splice product.

Figure 4.4. Suppression of non-canonical alternative 3´-splice sites in *PHO85*, *MAF1*, and *BET4* by Prp18p. A. Schematic diagrams of the *PHO85*, *MAF1*, and *BET4* transcripts illustrating the annotated (S) and alternative (*3´-AS) splicing events that occur for each transcript. The alternative 3´-splicing events were unambiguously confirmed by RT-PCR, cloning, and sequencing. Numeric values indicate the number of nucleotides in the coding sequences of the exons and do not include untranslated regions. Diagrams are not to scale. **B.** RT-PCR analysis of *PHO85* splicing in wild-type, NMD, and *prp18* mutant strains. Band labels indicate the unspliced (US) and spliced products as depicted in the diagrams in A. **C.** RT-PCR analysis of *MAF1* (upper panel) and *BET4* (lower panel) splicing in wild-type, NMD, and *prp18* mutant strains. Band labels are similar to panel B. The asterisk in the *BET4* panel denotes an unidentified splice product.

Figure 4.5. Mutations within various regions of Prp18p differentially impact alternative splicing. A. Growth of wild-type and *prp18*Δ*upf1*Δ strains complemented with various *PRP18* plasmids: empty pUG35 (vector) or plasmids expressing wild-type *PRP18*, a helix 2 mutant (*prp18-h2*), a helix 5 mutant (*prp18-h5*), or a mutant in which the conserved loop is completely deleted (*prp18*-ΔCRt). Strains were grown to exponential phase in –URA media, serially diluted in 3-fold spot dilutions onto –URA plates, and incubated at 30°C for the indicated lengths of time. **B.** RT-PCR analysis of *NYV1* splicing in the strains indicated in panel A. Band labels indicate unspliced (US), alternatively-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product.

Figure 4.6. Heat-sensitive *prp22* and cold-sensitive *prp43* mutants do not activate alternative 3'-SS. A. RT-PCR analysis of *NYV1* splicing in wild-type, *prp18*Δ, or heat-sensitive *prp22* mutant strains in NMD-positive or mutant backgrounds. All strains were grown to exponential phase in SDC media at 25°C and then shifted to 37°C for 1 hour. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product. **B.** RT-PCR analysis of *NYV1* splicing in *upf1*Δ NMD-deficient strains. Shown are wild-type, *prp18*Δ, *prp43* null mutant strains carrying pRS315 plasmids expressing wild-type *PRP43*, *prp43-Q423E*, or *prp22-Q423N* alleles. The wild-type and *prp18*Δ strains (lanes 1-2 and 6-7) were grown to exponential phase in SDC media while the strains carrying pRS315 plasmids (lanes 3-5 and 8-10) were grown in -LEU media. Bands are labeled similarly to panel A.

Figure 4.7. Alternative splicing in 3′-SS fidelity mutants of *PRP8* and *PRP22.* A. Analysis of *NYV1* splicing in WT, *prp18Δ*, or *PRP8* point mutant strains in either NMD-positive or mutant backgrounds. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and

spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product. **B.** RT-PCR analysis of *NYV1* (left panel) and *RAD14* (right panel) splicing in *upf1*Δ NMD-deficient strains. Shown are wild-type, *prp18*Δ, *prp22* null mutant strains carrying pRS315 plasmids expressing wild-type *PRP22*, *prp22-R805A*, or *prp22-G810A* alleles. The wild-type and *prp18*Δ strains (lanes 1-2) were grown to exponential phase in SDC media while the strains carrying pRS315 plasmids (lanes 3-5) were grown in -LEU media. Bands are labeled similarly to panel A above. Band labels for *RAD14* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts as well as two additional alternative 3´-spliced products labeled as x and y.

Figure 4.8. A shift to cold temperatures does not trigger non-canonical alternative splicing in cold-sensitive *PRP22* mutants. A. Spot dilution growth assay showing *prp22*Δ null mutant yeast in either a wild-type or *upf1*Δ NMD mutant background carrying pRS315 plasmids expressing either wild-type *PRP22*, *prp22-R805A*, or *prp22-G810A* alleles. Strains were grown to exponential phase in –LEU media then serially diluted in 3-fold spot dilutions onto –LEU plates and incubated at the indicated temperatures for 3 days. **B.** RT-PCR analysis of *NYV1* splicing in *upf1*Δ NMD-deficient strains. Shown are wild-type, *prp18*Δ, *prp22* null mutant strains carrying pRS315 plasmids expressing wild-type *PRP22*, *prp22-R805A*, or *prp22-G810A* alleles. The wild-type and *prp18*Δ strains (lanes 1-2, 6-7, 11-12) were grown to exponential phase in SDC media while the strains carrying pRS315 plasmids (lanes 3-5, 8-10, 13-15) were grown in -LEU media at 30°C and then shifted to either 25°C or 16°C for 2 hours. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product.

Figure 4.9. Mutating the interface between Prp18p and Prp8p activates alternative 3′-SS selection. A. Pymol view of the yeast catalytically active spliceosome showing interactions

between residues P-1984 and L-1988 of Prp8p and residues Q-181 and N-190 of Prp18p. Structure image adapted from (Yan et al., 2016). **B.** RT-PCR analysis of *NYV1* splicing in wild-type, *prp18*Δ, or *PRP8* point mutant strains in NMD-positive or deficient backgrounds. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product. **C.** Spot dilution growth assay showing the wild-type strain with an empty pUG35 vector or *prp18*Δ*upf1*Δ strains complemented with empty pUG35 or pUG35 vectors expressing wild-type or mutant *PRP18* alleles. Strains were grown to exponential phase in –URA media, serially diluted in 3-fold spot dilutions onto –URA plates and incubated at 30°C for 4 days. **D.** RT-PCR analysis of *NYV1* splicing in the strains shown in panel C. Bands are labeled similarly to panel B.

Figure 4.10. Overexpression of Slu7p partially rescues growth defect of *prp18*Δ but does not suppress non-canonical alternative splicing for *NYV1*. A. Spot dilution assay for WT strain with empty multicopy YEp24 plasmid and *prp18*Δ*upf1*Δ strains carrying empty YEp24 or YEp24 expressing the wild-type *PRP18* or *SLU7* alleles. Strains were grown to exponential phase in –URA media, serially diluted in 3-fold spot dilutions onto –URA plates, and incubated at 30°C for 3 days. **B**. RT-PCR analysis of *NYV1* splicing in WT, *upf1*Δ, *prp18*Δ, or *upf1*Δ*prp18*Δ strains carrying an empty multicopy YEp24 plasmid or YEp24 expressing wild-type *SLU7*. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product.

Figure 4.11. Splicing analysis for *RPL22B* and *snR17* in strains overexpressing *SLU7*. A. RT-PCR analysis of *RPL22B* in WT, *upf1*Δ, *prp18*Δ, or *upf1*Δ*prp18*Δ strains carrying an empty multicopy YEp24 plasmid or YEp24 expressing *SLU7*. Band labels for *NYV1* indicate unspliced (US), alternative-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified

NMD-sensitive splice product. Band labels for *RPL22B* indicate unspliced (US), alternative 5'-spliced (*AS 5'), and spliced (S) transcripts. **B.** RT-PCR analysis of *snR17A* and *snR17B* unspliced (US) and spliced (S) transcripts in wild-type and *prp18*Δ strains with empty YEp24 or YEp24 expressing *SLU7*. PRP5-FRB and PRP28-FRB samples in which Prp5p and Prp28p were depleted from the nucleus by the anchor away technique were used as splicing-deficient negative controls, with cDNA generated by either a gene-specific RT primer or a random hexamer.

Figure 4.12. Nuclear depletion of Prp18p, but not Slu7p, activates the alternative 3-splice site in *NYV1*. A. RT-PCR analysis of *NYV1* splicing in strains containing the FRB tag on Prp18p or Slu7p to enable nuclear depletion upon addition of rapamycin. Cells were grown in YPD media to exponential phase, after which rapamycin was added to a final concentration of 1 µg/ml and cells were grown for an additional 120 mins. Band labels indicate unspliced (US), alternatively-spliced (*AS 3´), and spliced (S) transcripts. **B.** RT-PCR analysis of *NYV1* in Prp18p and Slu7p anchor away strains as in panel A, but with the inclusion of *upf1*△ NMD mutants a control strain lacking an FRB tag and a shift time of 60 minutes. Bands are labeled similarly to panel A. The asterisk denotes an unidentified NMD-sensitive splice product.

Figure 4.13. A *slu7* mutation that disrupts Prp18p-Slu7p interaction partially activates the *NYV1* alternative 3´-splice site. A. Spot dilution assay for *slu7*Δ null mutant strains rescued with plasmids expressing wild-type and mutant *SLU7* alleles. Strains were grown to exponential phase in –TRP media, serially diluted in 3-fold spot dilutions onto –TRP plates, and incubated at 30 and 37°C for 2 days. **B.** RT-PCR analysis of *NYV1* splicing in *upf1*Δ mutant strains carrying the plasmid-based *SLU7* alleles described in panel A as well as the *prp18*Δ null mutant. Strains carrying the *SLU7* alleles (lanes 1-4, 6-9) were grown in –TRP media to exponential phase at 30°C and then shifted to 37°C for 1 hour. The *prp18Δupf1Δ* strain (lanes 5 and 10) was grown in

SDC media and treated similarly. Band labels indicate unspliced (US), alternatively-spliced (*AS 3´), and spliced (S) transcripts. The asterisk denotes an unidentified NMD-sensitive splice product.

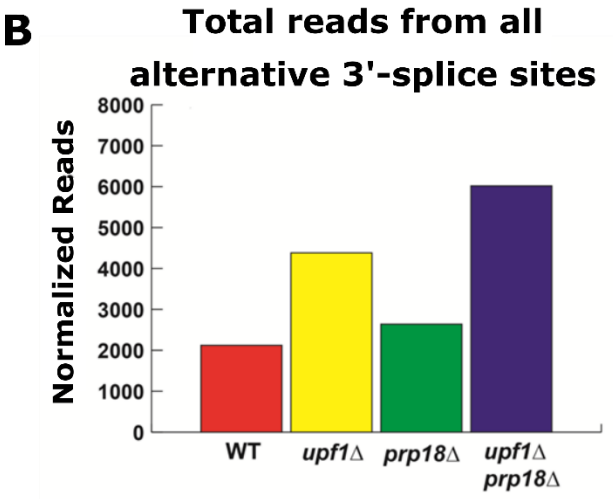
Figure 4.1

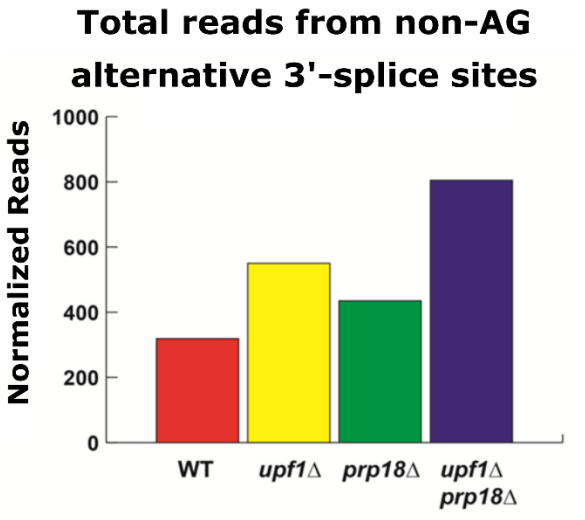
A

RNA-Seq Read information

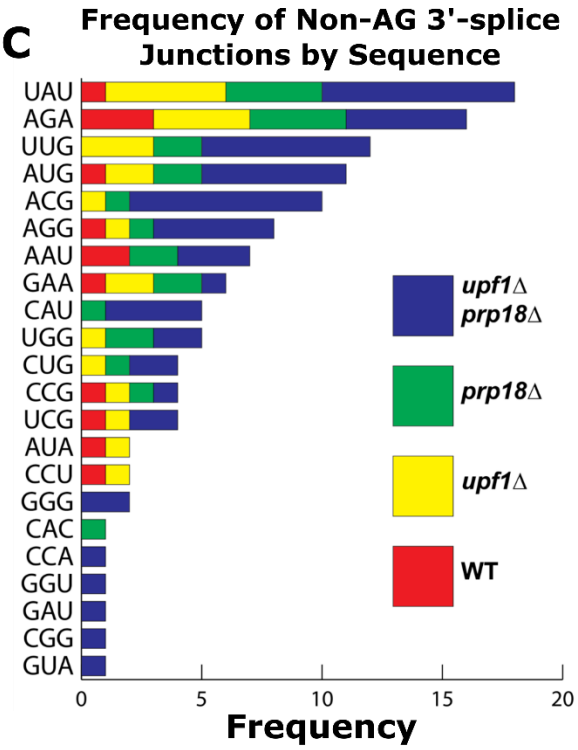
Strain	Uniquely Mapped Reads	Reads Mapping to Any Splice Junction	% of Unique Reads Mapping to Splice Junctions
WT	45,351,821	2,020,071	4.5%
upf1Δ	43,815,382	1,619,269	3.7%
prp18Δ	52,030,754	1,745,469	3.4%
prp18Δupf1Δ	47,077,091	1,521,510	3.2%

B





C



D

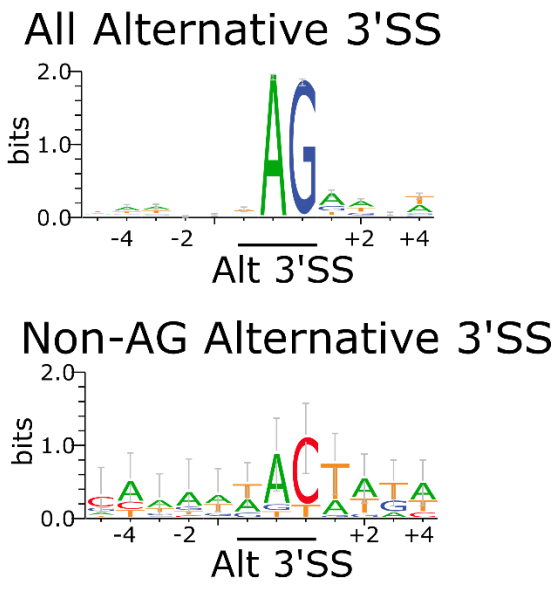
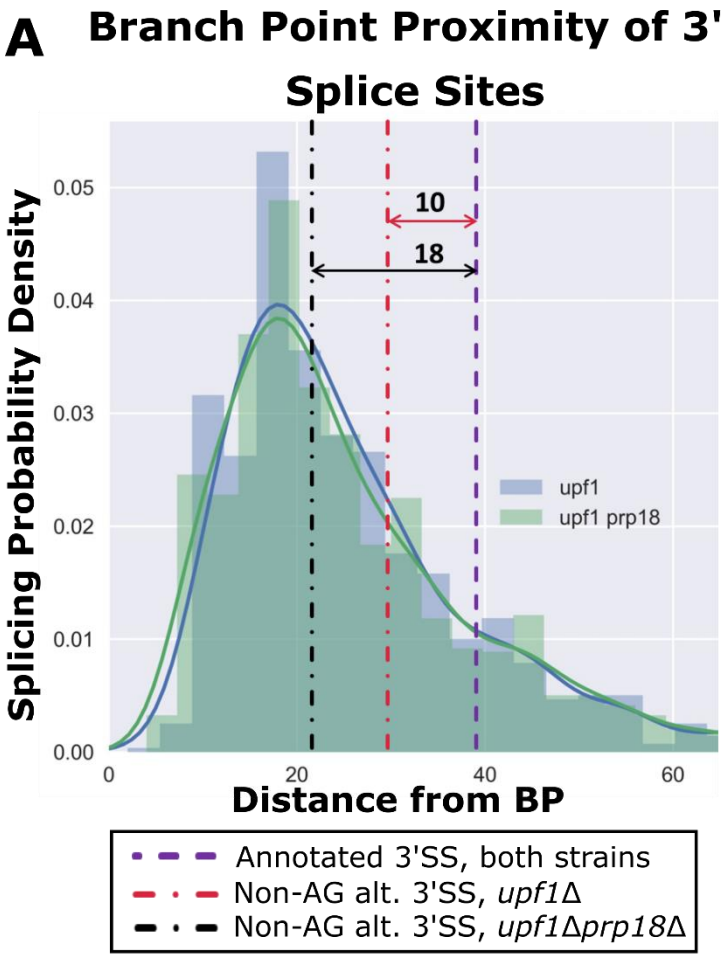


Figure 4.2

A



B

	Intron-containing RPGs	Intron-containing non-RPGs	Total
# non-AG alternative 3′-splice sites	243(58.9%)	170(41.1%)	413(100%)
# non-AG alternative 3′-splice sitescomprising ≥1% of all splicing	27(32.1%)	57(67.9%)	84(100%)
All intron-containing genes(Ares et al., 1999)	95 (41.5%)	134 (58.5%)	229 (100%)

Figure 4.3

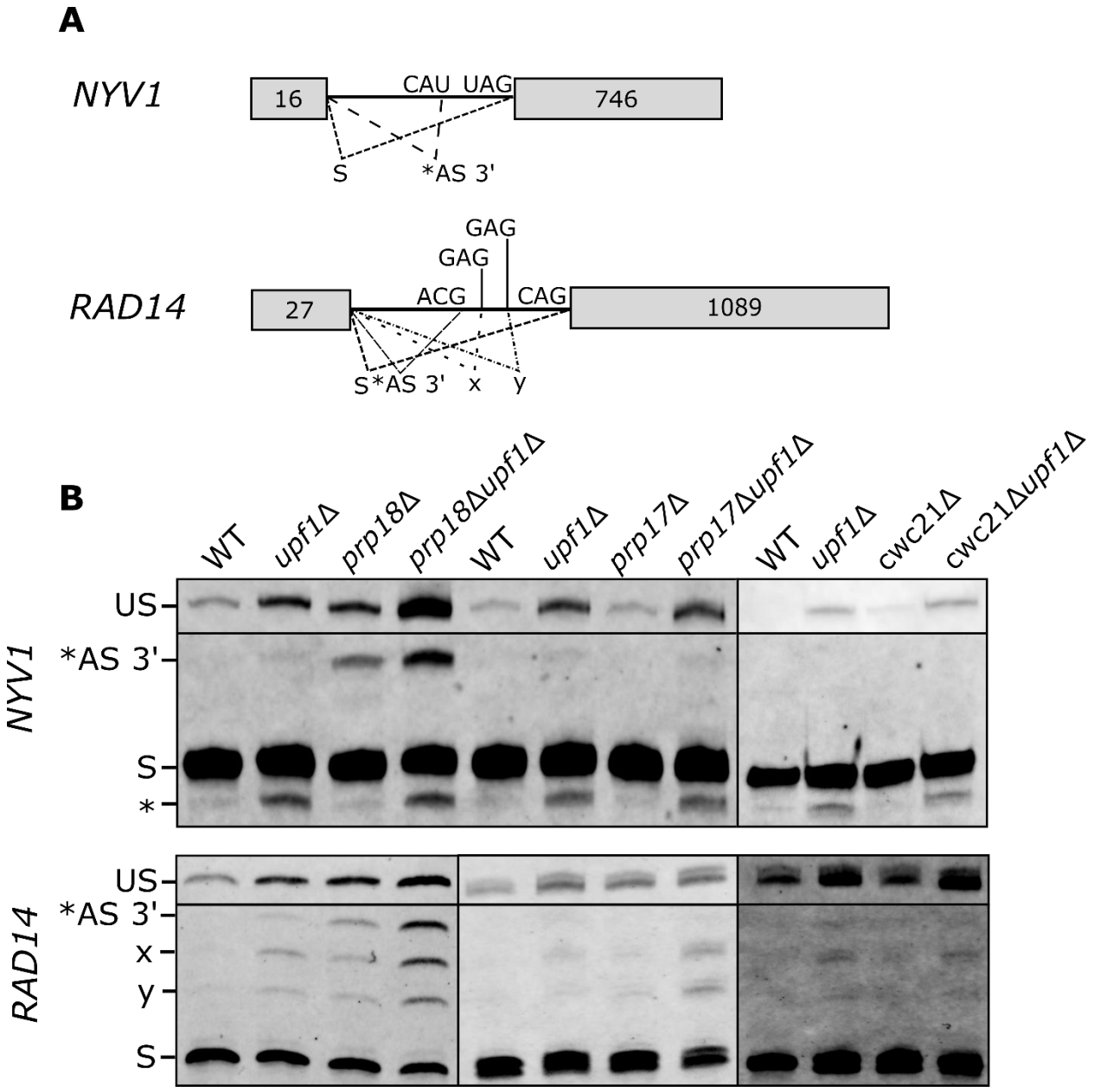

A
NYV1
16
CAU UAG
746
S
*AS 3'
RAD14
GAG
GAG
ACG
CAG
27
1089
S*AS 3'
x
y
B
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
WT
upf1Δ
prp17Δ
prp17Δupf1Δ
WT
upf1Δ
cwc21Δ
cwc21Δupf1Δ
NYV1
US
*AS 3'
S
*
RAD14
US
*AS 3'
x
y
S

Figure 4.4

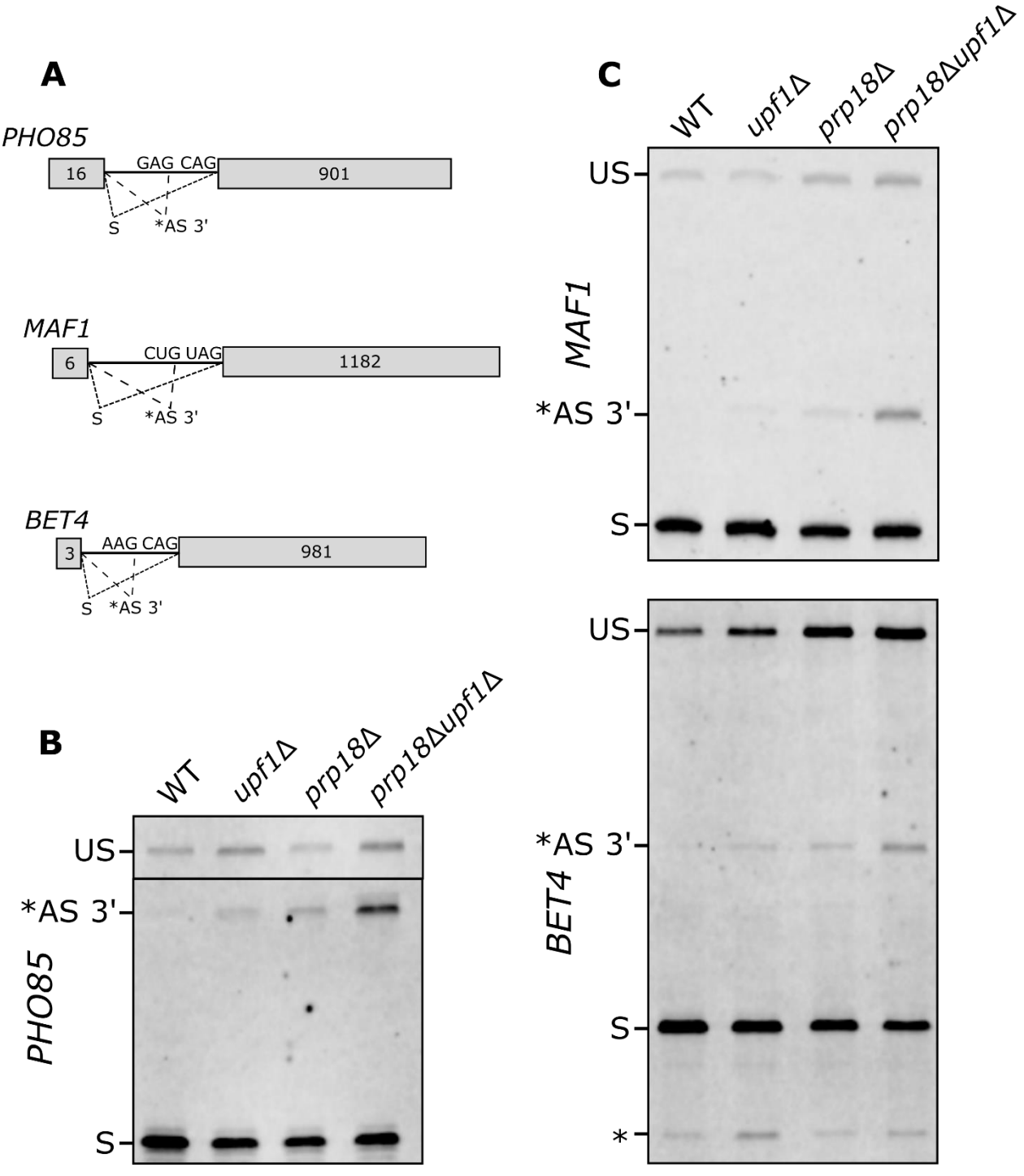

A
PHO85
16
GAG CAG
901
S
*AS 3'
MAF1
6
CUG UAG
1182
S
*AS 3'
BET4
3
AAG CAG
981
S
*AS 3'
B
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
US
*AS 3'
PHO85
S
C
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
US
MAF1
*AS 3'
S
US
*AS 3'
BET4
S
*

Figure 4.5

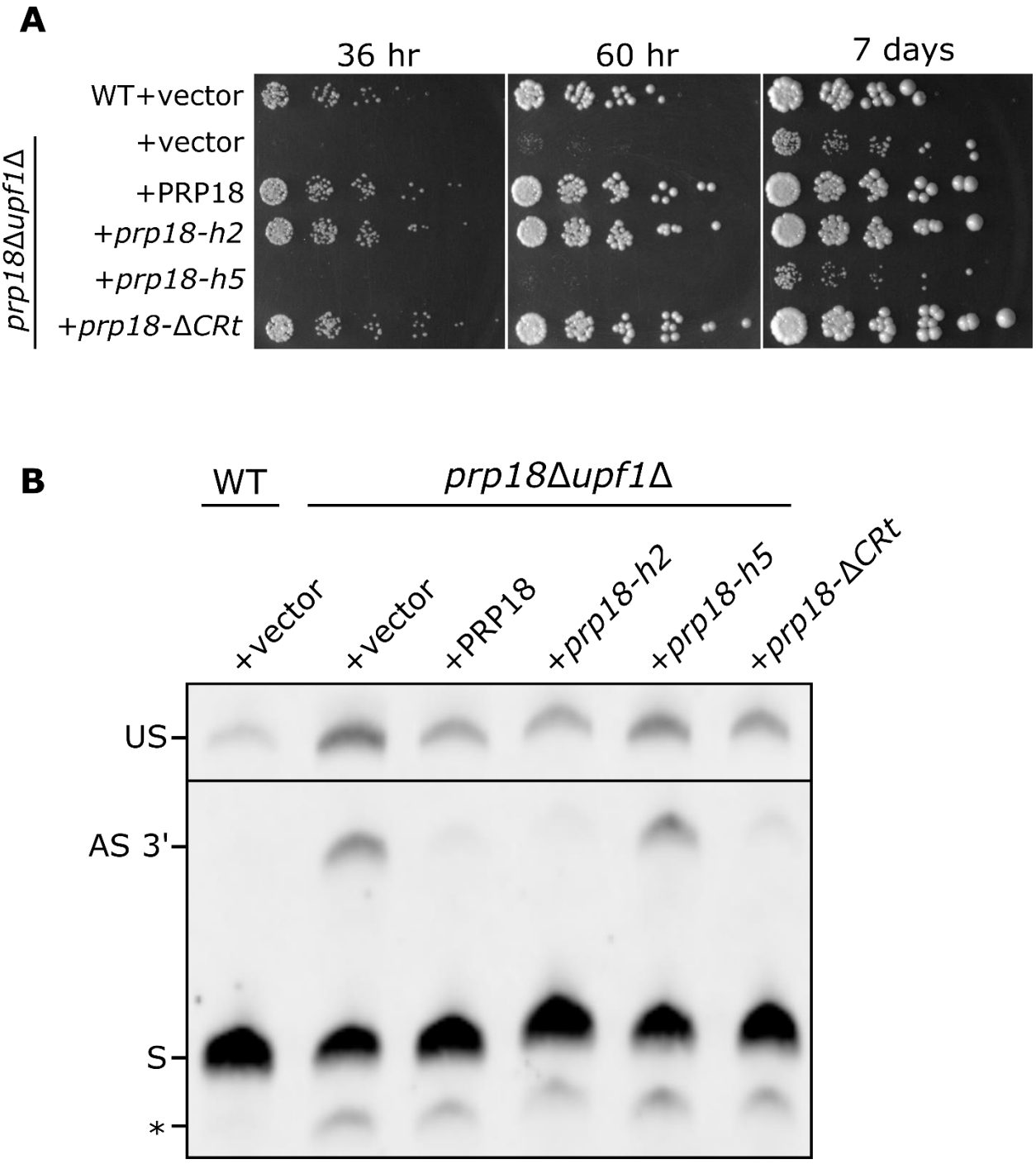

A
36 hr
60 hr
7 days
WT+vector
+vector
+PRP18
+prp18-h2
+prp18-h5
+prp18-ΔCRt
prp18Δupf1Δ
B
WT
prp18Δupf1Δ
+vector
+vector
+PRP18
+prp18-h2
+prp18-h5
+prp18-ΔCRt
US
AS 3'
S
*

Figure 4.6

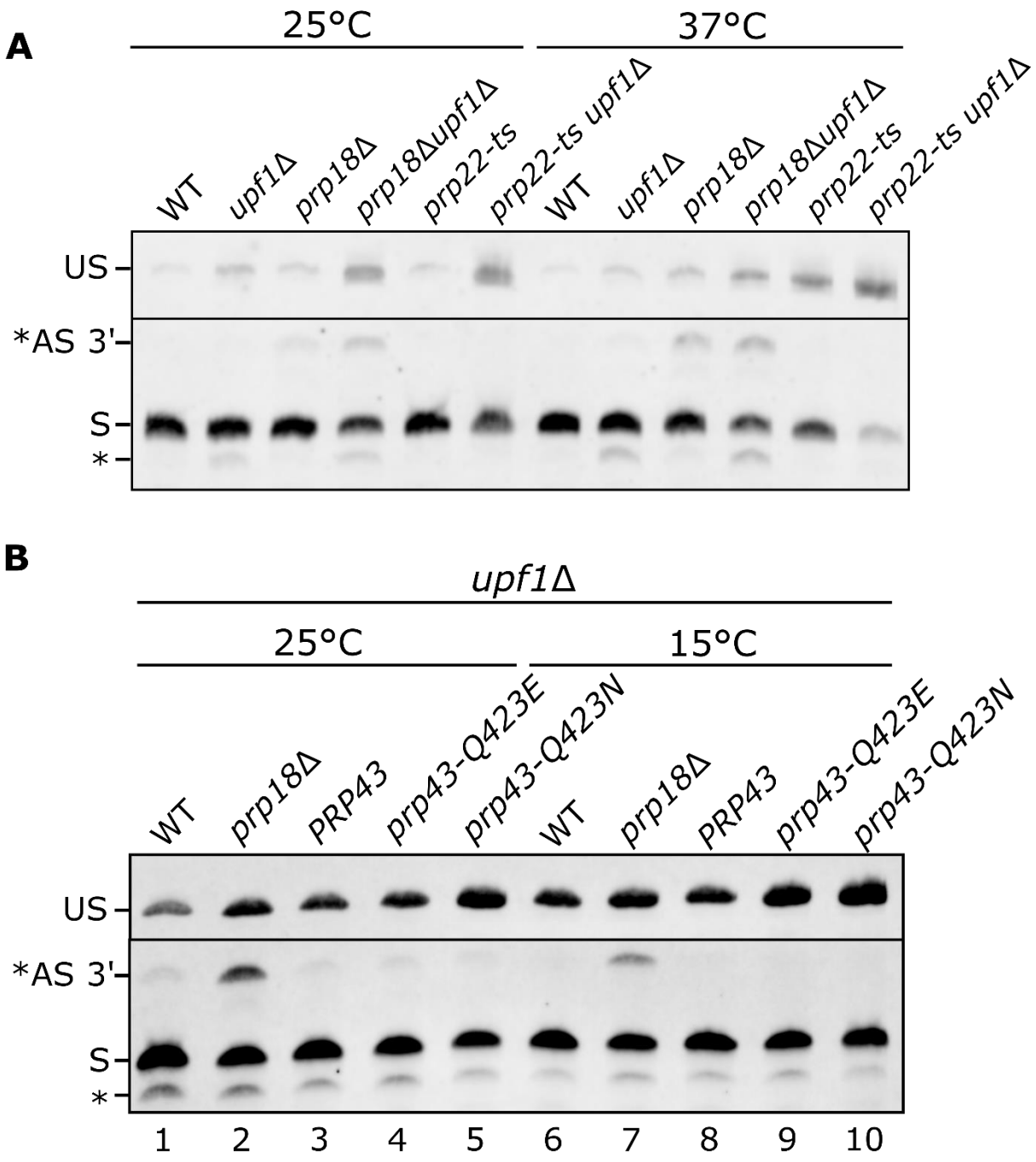

A
25°C
37°C
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
prp22-ts
prp22-ts upf1Δ
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
prp22-ts
prp22-ts upf1Δ
US
*AS 3'
S
*
B
upf1Δ
25°C
15°C
WT
prp18Δ
PRP43
prp43-Q423E
prp43-Q423N
WT
prp18Δ
PRP43
prp43-Q423E
prp43-Q423N
US
*AS 3'
S
*
1
2
3
4
5
6
7
8
9
10

Figure 4.7

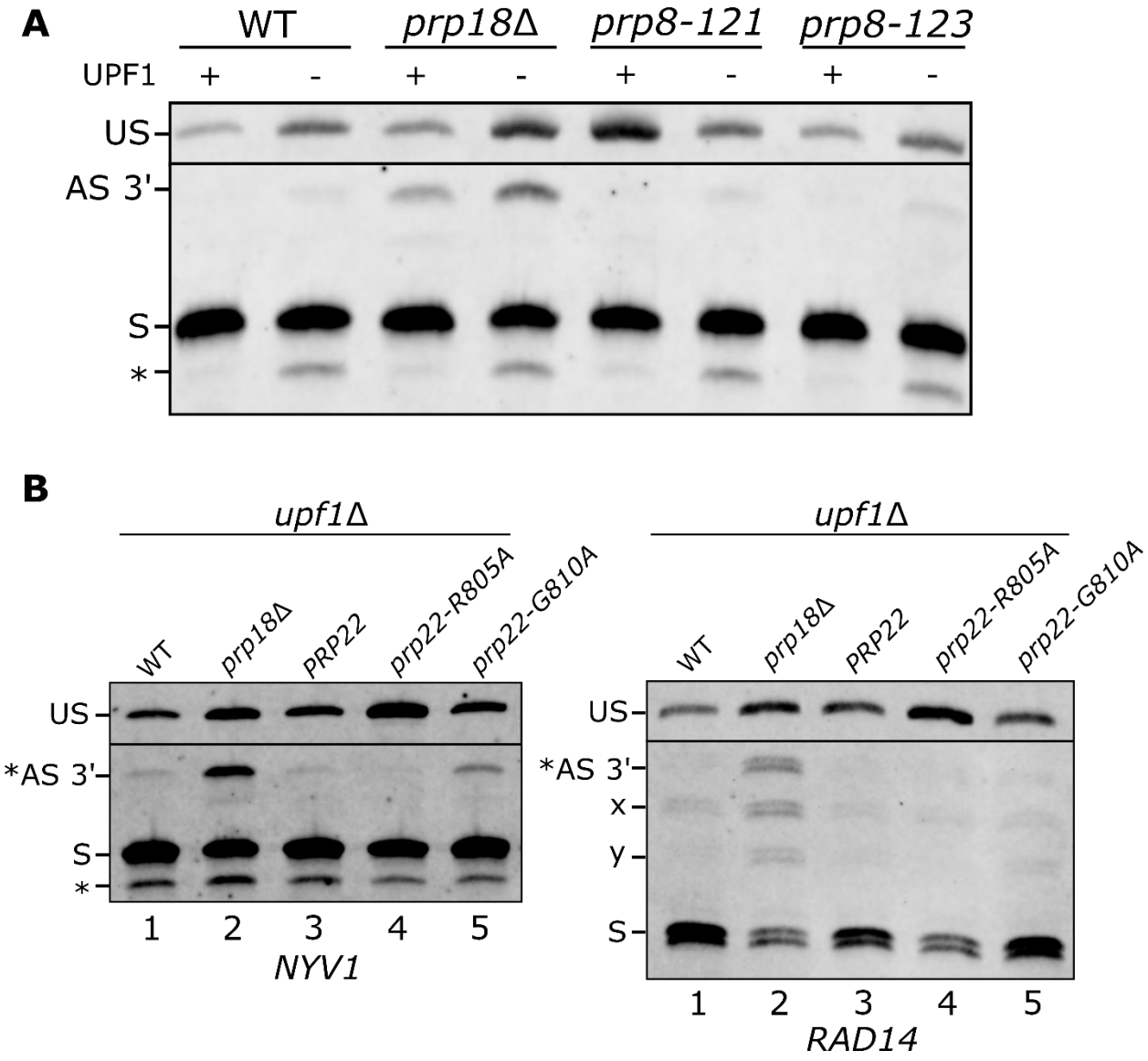

A
WT
prp18Δ
prp8-121
prp8-123
UPF1
+
-
+
-
+
-
+
-
US
AS 3'
S
*
B
upf1Δ
upf1Δ
WT
prp18Δ
PRP22
prp22-R805A
prp22-G810A
US
*AS 3'
S
*
1
2
3
4
5
NYV1
x
y
RAD14

Figure 4.8

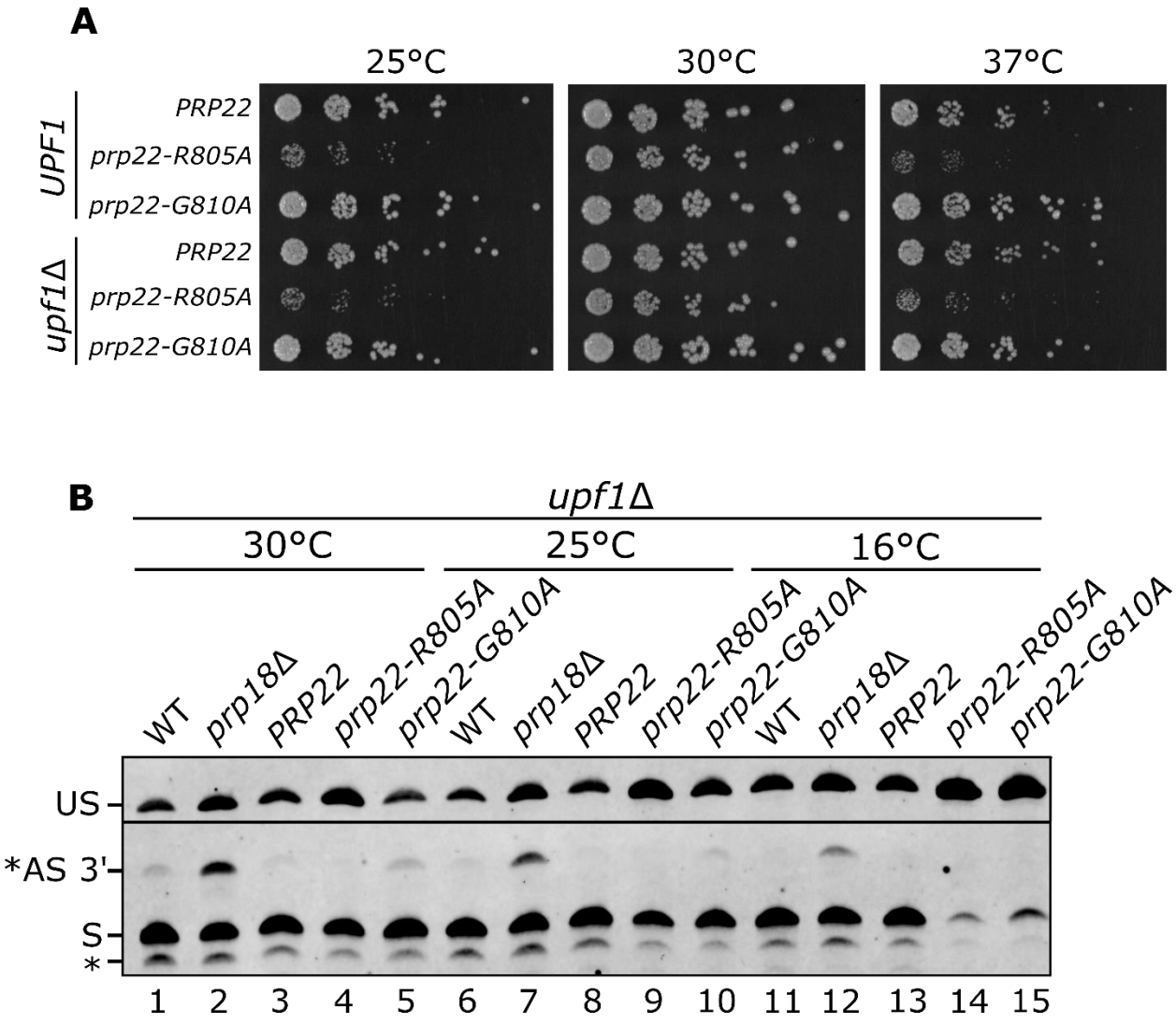

A
25°C
30°C
37°C
UPF1
PRP22
prp22-R805A
prp22-G810A
upf1Δ
PRP22
prp22-R805A
prp22-G810A
B
upf1Δ
30°C
25°C
16°C
WT
prp18Δ
PRP22
prp22-R805A
prp22-G810A
US
*AS 3'
S
*
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Figure 4.9

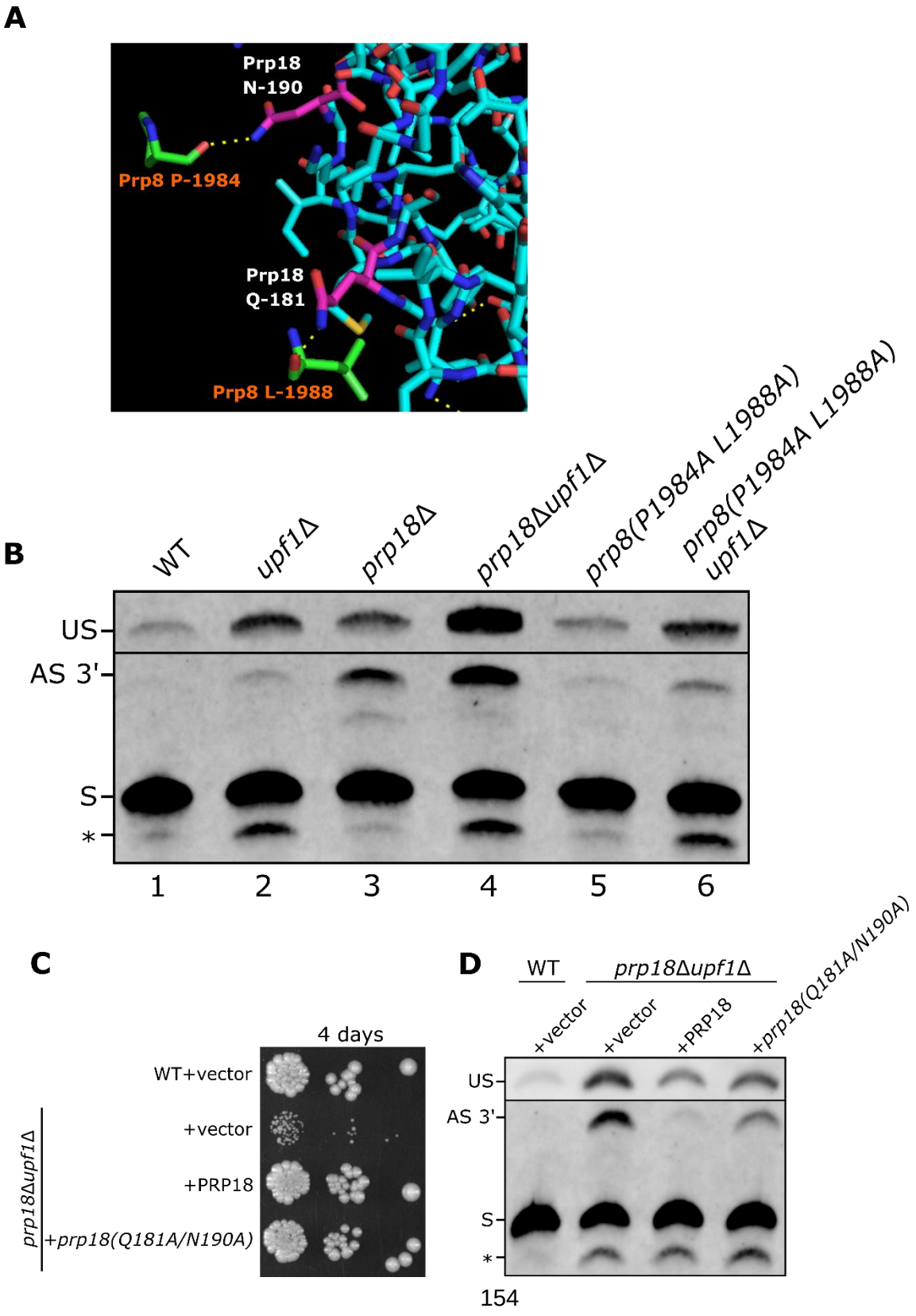

A
Prp18
N-190
Prp8 P-1984
Prp18
Q-181
Prp8 L-1988
B
WT
upf1Δ
prp18Δ
prp18Δupf1Δ
prp8(P1984A L1988A)
prp8(P1984A L1988A)
upf1Δ
US
AS 3'
S
*
1
2
3
4
5
6
C
4 days
WT+vector
+vector
+PRP18
+prp18(Q181A/N190A)
prp18Δupf1Δ
D
WT
prp18Δupf1Δ
+vector
+vector
+PRP18
+prp18(Q181A/N190A)
US
AS 3'
S
*
154

Figure 4.10

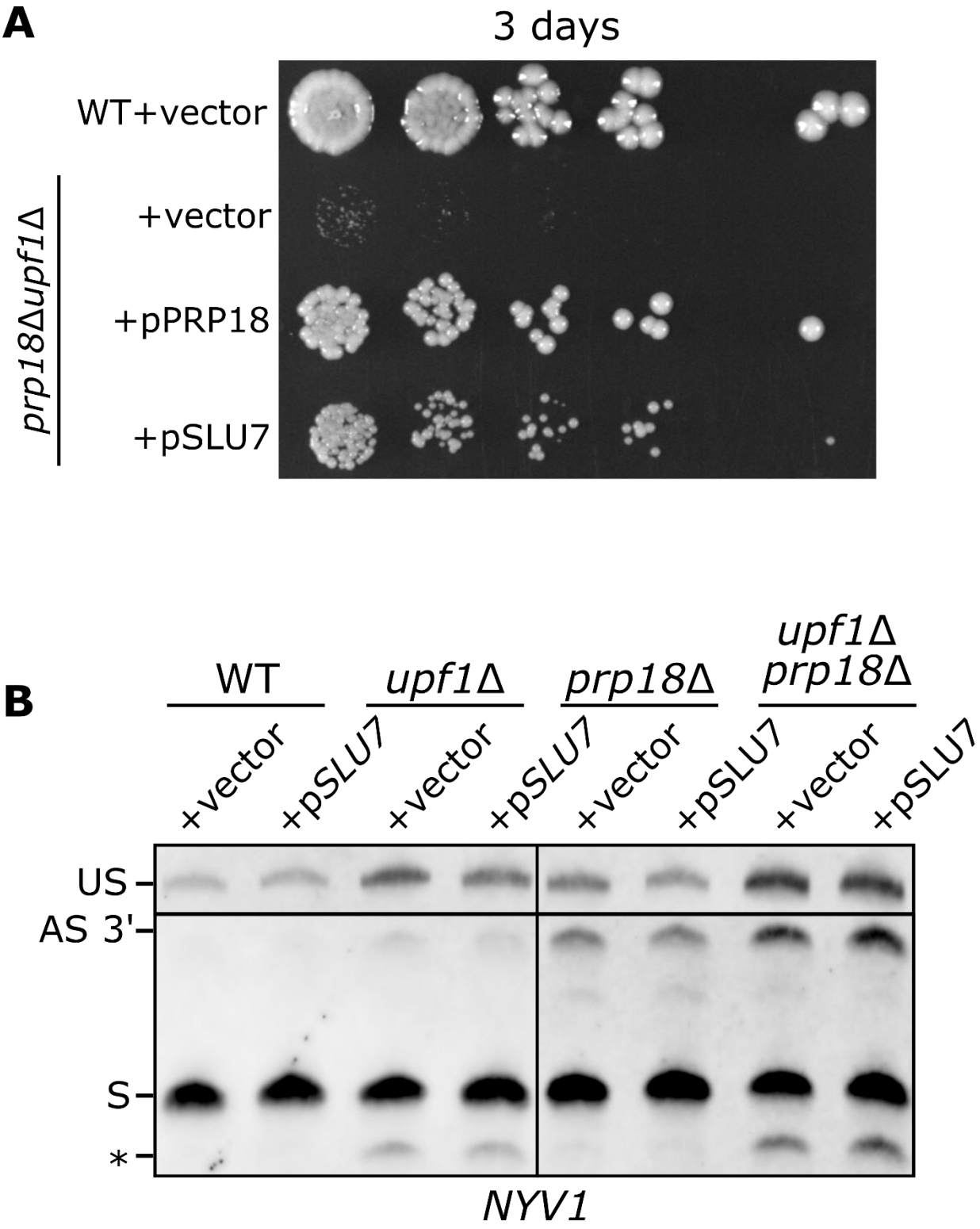

A
3 days
WT+vector
prp18Δupf1Δ
+vector
+pPRP18
+pSLU7
B
WT
upf1Δ
prp18Δ
upf1Δ prp18Δ
+vector
+pSLU7
+vector
+pSLU7
+vector
+pSLU7
+vector
+pSLU7
US
AS 3'
S
*
NYV1

Figure 4.11

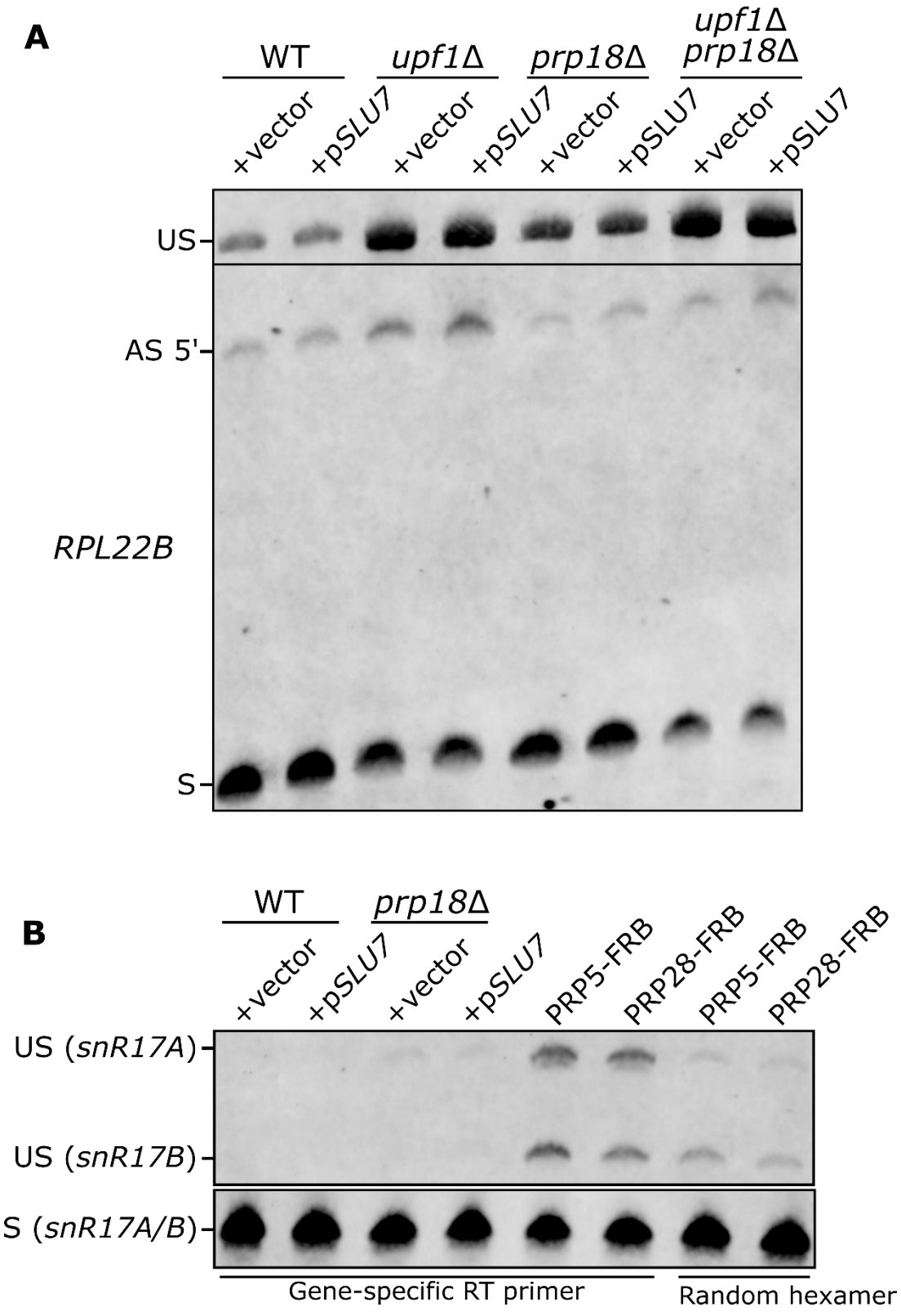

A
WT
upf1Δ
prp18Δ
upf1Δ prp18Δ
+vector
+pSLU7
US
AS 5'
RPL22B
S
B
WT
prp18Δ
+vector
+pSLU7
PRP5-FRB
PRP28-FRB
US (snR17A)
US (snR17B)
S (snR17A/B)
Gene-specific RT primer
Random hexamer

Figure 4.12

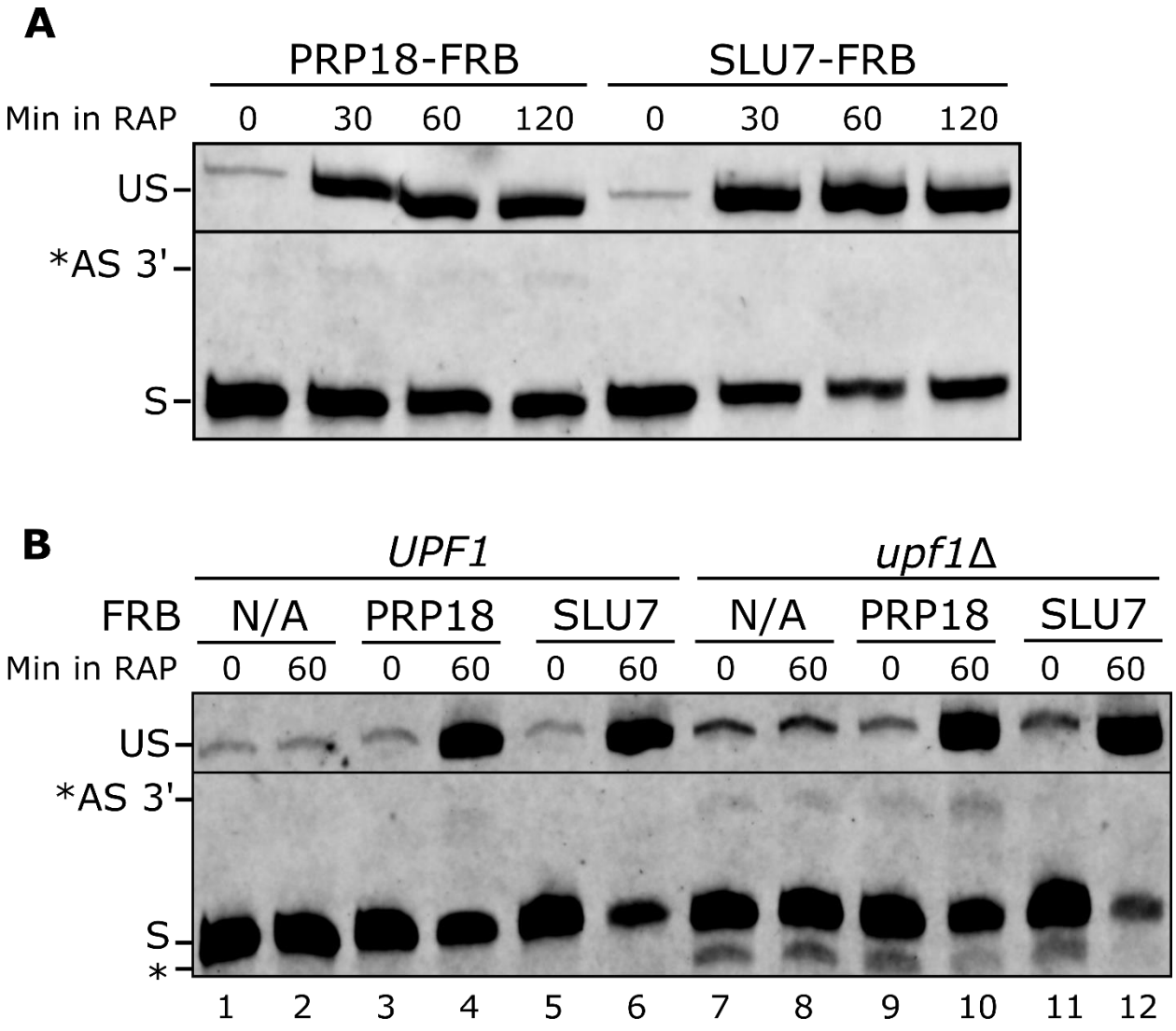
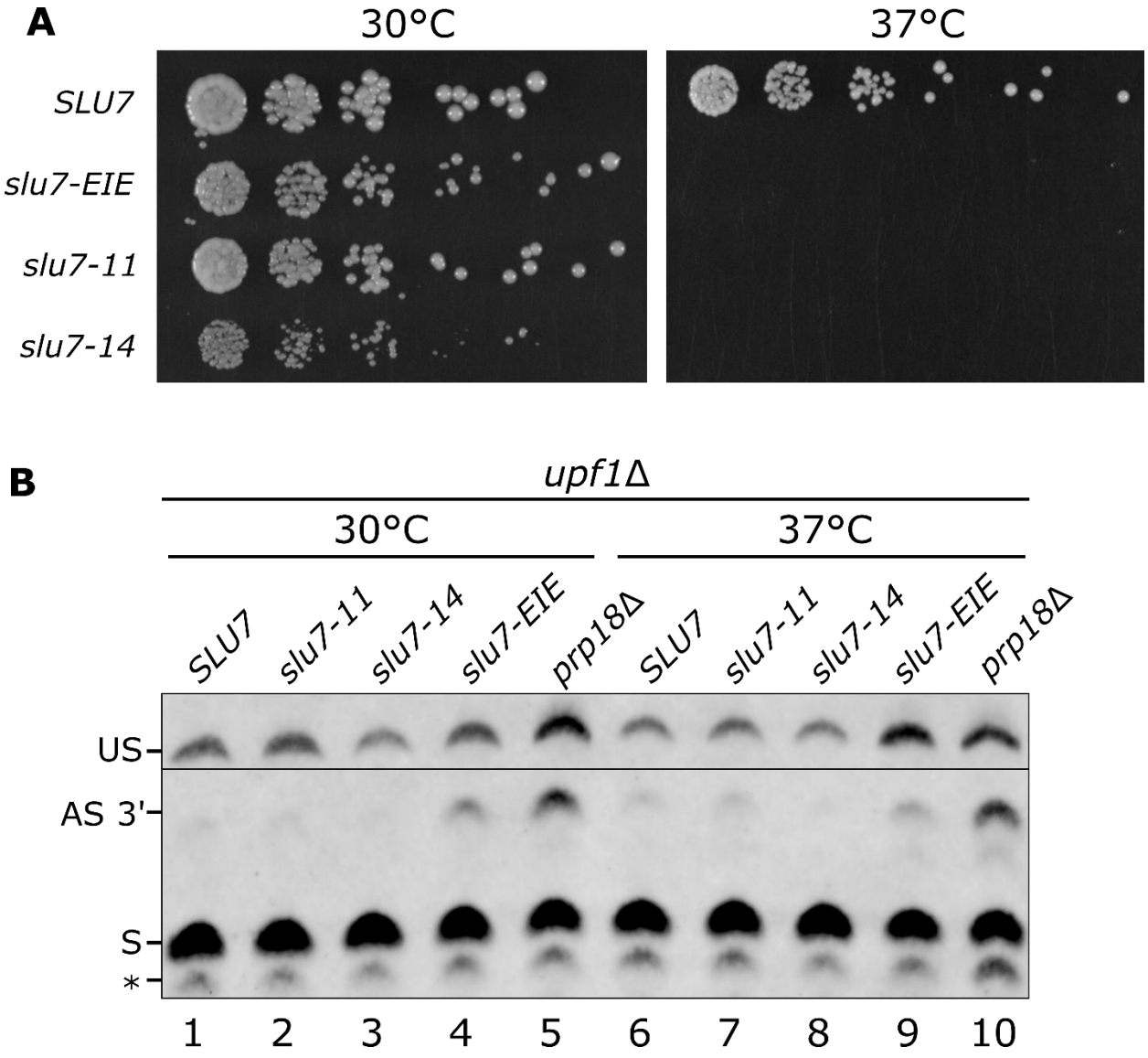

A
PRP18-FRB
SLU7-FRB
Min in RAP
0
30
60
120
0
30
60
120
US
*AS 3'
S
B
UPF1
upf1Δ
FRB
N/A
PRP18
SLU7
N/A
PRP18
SLU7
Min in RAP
0
60
0
60
0
60
0
60
0
60
0
60
US
*AS 3'
S
*
1
2
3
4
5
6
7
8
9
10
11
12

Figure 4.13


A
30°C
37°C
SLU7
slu7-EIE
slu7-11
slu7-14
B
upf1Δ
30°C
37°C
SLU7
slu7-11
slu7-14
slu7-EIE
prp18Δ
SLU7
slu7-11
slu7-14
slu7-EIE
prp18Δ
US
AS 3'
S
*
1
2
3
4
5
6
7
8
9
10

References

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.

Aronova, A., Dagmar, B., Crotti, L.B., Horowitz, D.S., and Schwer, B. (2007). Functional interactions between Prp8, Prp18, Slu7, and U5snRNA during the second step of pre-mRNA splicing. RNA *13*, 1437–1444.

Bacíková, D., and Horowitz, D.S. (2002). Mutational analysis identifies two separable roles of the Saccharomyces cerevisiae splicing factor Prp18. RNA *8*, 1280–1293.

Bacíková, D., and Horowitz, D.S. (2005). Genetic and Functional Interaction of Evolutionarily Conserved Regions of the Prp18 Protein and the U5 snRNA. Mol Cell Biol *25*, 2107–2116.

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner.

Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: A Sequence Logo Generator. Genome Res. *14*, 1188–1190.

Crotti, L.B., and Horowitz, D.S. (2009). Exon sequences at the splice junctions affect splicing fidelity and alternative splicing. PNAS *106*, 18954–18959.

Crotti, L.B., Bačíková, D., and Horowitz, D.S. (2007). The Prp18 protein stabilizes the interaction of both exons with the U5 snRNA during the second step of pre-mRNA splicing. Genes Dev. *21*, 1204–1216.

Dicarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343.

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

Fica, S.M., Oubridge, C., Galej, W.P., Wilkinson, M.E., Bai, X.-C., Newman, A.J., and Nagai, K. (2017). Structure of a spliceosome remodelled for exon ligation. Nature *542*, 377–380.

Fourmann, J.B., Tauchert, M.J., Ficner, R., Fabrizio, P., and Lührmann, R. (2017). Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. Nucleic Acids Res. *45*, 4068–4080.

Frank, D., and Guthrie, C. (1992). An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. Genes Dev. *6*, 2112–2124.

Gabunilas, J., and Chanfreau, G. (2016). Splicing-Mediated Autoregulation Modulates Rpl22p Expression in Saccharomyces cerevisiae. PLoS Genet. *12*.

Gautam, A., Grainger, R.J., Vilardell, J., Barrass, J.D., and Beggs, J.D. (2015). Cwc21p promotes the second step conformation of the spliceosome and modulates 3′ splice site selection. Nucleic Acids Res. *43*, 3309–3317.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gillespie, A., Gabunilas, J., Jen, J.C., and Chanfreau, G.F. (2017). Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in S. cerevisiae. RNA *23*, 466–472.

Gould, G.M., Paggi, J.M., Guo, Y., Phizicky, D. V., Zinshteyn, B., Wang, E.T., Gilbert, W. V., Gifford, D.K., and Burge, C.B. (2016). Identification of new branch points and unconventional introns in Saccharomyces cerevisiae. RNA *22*, 1522–1534.

Grainger, R.J. (2005). Prp8 protein: At the heart of the spliceosome. RNA *11*, 533–557.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The Anchor-Away Technique: Rapid, Conditional Establishment of Yeast Mutant Phenotypes. Mol. Cell *31*, 925–932.

Horowitz, D.S. (2012). The mechanism of the second step of pre-mRNA splicing. Wiley Interdiscip. Rev. RNA *3*, 331–350.

Horowitz, D.S., and Abelson, J. (1993). A U5 Small Nuclear Ribonucleoprotein Particle Protein Involved Only in the 2nd Step of Pre-Messenger RNA Splicing in Saccharomyces-Cerevisiae. Mol Cell Biol *13*, 2959–2970.

James, S.A., Turner, W., and Schwer, B. (2002). How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. RNA *8*, 1068–1077.

Jiang, J., Horowitz, D.S., and Xu, R.M. (2000). Crystal structure of the functional domain of the splicing factor Prp18. PNAS *97*, 3022–3027.

Kawashima, T., Douglass, S., Gabunilas, J., Pellegrini, M., and Chanfreau, G.F. (2014). Widespread Use of Non-productive Alternative Splice Sites in Saccharomyces cerevisiae. PLoS Genet. *10*.

Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A., and Murphy, J.W. (2004). Introns and splicing elements of five diverse fungi. Eukaryot. Cell *3*, 1088–1100.

Ma, P., and Xia, X. (2011). Factors affecting splicing strength of yeast genes. Comp. Funct. Genomics *2011*.

Maquat, L.E. (2004). Nonsense-Mediated mRNA Decay: Splicing, Translation And mRNP Dynamics. Nat. Rev. Mol. Cell Biol. *5*, 89–99.

Mayas, R.M., Maita, H., and Staley, J.P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat. Struct. Mol. Biol. *13*, 482–490.

Mayas, R.M., Maita, H., Semlow, D.R., and Staley, J.P. (2010). Spliceosome discards intermediates via the DEAH box ATPase Prp43p. PNAS *107*, 10020–10025.

O'Keefe, R.T., and Newman, A.J. (1998). Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. EMBO J. *17*, 565–574.

Ohrt, T., Odenwälder, P., Dannenberg, J., Prior, M., Warkocki, Z., Schmitzová, J., Karaduman, R., Gregor, I., Enderlein, J., Fabrizio, P., et al. (2013). Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. RNA *19*, 902–915.

Parker, R., and Siliciano, P.G. (1993). Evidence for an Essential Non-Watson-Crick Interaction Between the 1st and Last Nucleotides of a Nuclear Pre-Messenger RNA Intron. Nature *361*, 660–662.

Roy, K., and Chanfreau, G. (2014). Stress-Induced Nuclear RNA Degradation Pathways Regulate Yeast Bromodomain Factor 2 to Promote Cell Survival. PLoS Genet. *10*.

Roy, K., Gabunilas, J., Gillespie, A., Ngo, D., and Chanfreau, G.F. (2016). Common genomic elements promote transcriptional and DNA replication roadblocks. Genome Res. *26*, 1363–1375.

Sapra, A.K., Arava, Y., Khandelia, P., and Vijayraghavan, U. (2004). Genome-wide analysis of Pre-mRNA splicing: Intron features govern the requirement for the second-step factor, Prp17 in Saccharomyces cerevisiae and Schizosaccharomyces pombe. J. Biol. Chem. *279*, 52437–52446.

Schreiber, K., Csaba, G., Haslbeck, M., and Zimmer, R. (2015). Alternative Splicing in Next Generation Sequencing Data of Saccharomyces cerevisiae. PLoS One *10*.

Schwer, B., and Gross, C.H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. EMBO J. *17*, 2086–2094.

Schwer, B., and Guthrie, C. (1992). A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. EMBO J. *11*, 5033–5039.

Schwer, B., and Meszaros, T. (2000). RNA helicase dynamics in pre-mRNA splicing. EMBO J. *19*, 6582–6591.

Semlow, D.R., Blanco, M.R., Walter, N.G., and Staley, J.P. (2016). Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. Cell *164*, 985–998.

Tsai, R.T., Fu, R.H., Yeh, F.L., Tseng, C.K., Lin, Y.C., Huang, Y.H., and Cheng, S.C. (2005). Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes Dev. *19*, 2991–3003.

Umen, J.G., and Guthrie, C. (1996). Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3??? splice site selection. Genetics *143*, 723–739.

Volanakis, A., Passoni, M., Hector, R.D., Shah, S., Kilchert, C., Granneman, S., and Vasiljeva, L. (2013). Spliceosome-mediated decay (SMD) regulates expression of nonintronic genes in budding yeast. Genes Dev. *27*, 2025–2038.

Wu, Q., and Krainer, A.R. (1999). AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. Mol Cell Biol *19*, 3225–3236.

Yan, C., Wan, R., Bai, R., Huang, G., and Shi, Y. (2016). Structure of a yeast step II catalytically activated spliceosome. Science (80-.). *355*, 149–155.

Zhang, X., and Schwer, B. (1997). Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. Nucleic Acids Res. *25*, 2146–2152.