Title

Quality Control of Splicing by the Nonsense-Mediated mRNA Decay Pathway in Saccharomyces cerevisiae

Permalink

https://escholarship.org/uc/item/75s1k198

Author

KAWASHIMA, TADASHI RYAN

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Quality Control of Splicing by the Nonsense-Mediated mRNA Decay Pathway

in *Saccharomyces cerevisiae*

A dissertation submitted in partial satisfaction of the

requirements for the Degree Doctor of Philosophy

in Biochemistry and Molecular Biology

by

Tadashi Ryan Kawashima

2013

ABSTRACT OF THE DISSERTATION

Quality Control of Splicing by the Nonsense-Mediated Decay Pathway

in *Saccharomyces cerevisiae*

by

Tadashi Ryan Kawashima

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2013

Professor Guillaume F. Chanfreau, Chair

Using the budding yeast, *Saccharomyces cerevisiae* as our model system we set out to investigate the role of nonsense-mediated mRNA decay (NMD) pathway in the quality control of premature translation termination codon (PTC) containing transcripts to include those resulting from splicing factor mutations as well as unproductive alternative splicing events. The function of many splicing factors in pre-mRNA splicing and their involvement in the processing of a specific subset of transcripts has often been defined through loss of function analysis. We show that NMD can mask some of the effects of splicing factor mutations, and the full role of the splicing factor cannot be understood unless the RNA degradation system that degrades unspliced precursors are also inactivated. Tiling microarrays showed that inactivation of the NMD factor Upf1p in combination with splicing factor mutants *prp17Δ* and *prp18Δ* resulted in a larger spectrum of splicing defects than in the single mutants. Analysis of these double mutants also hinted to the possibility of non-productive alternative splicing. This lead us to seek out and identify splicing defects arising from alternative splice site selection through RNA-Sequencing

analysis of wild-type and NMD mutant (*upf1Δ, upf2Δ,* and *upf3Δ*) strains. We found that a large fraction of intron containing genes exhibit alternative splicing, but are masked by NMD because they generate PTCs. Analysis of splicing factor mutations combined with *upf1Δ* revealed the role of specific splicing factors in governing the use of these alternative splice sites. Furthermore, we show the use of a non-productive alternative 5' splice site in *RPL22B* to be regulated during conditions of stress. These studies bring to light an unexpected flexibility of the spliceosome in splice site selection and the role of NMD in limiting the accumulation of these erroneous transcripts. Finally, we identified a small subset of two intron containing genes whose erroneous transcripts are not targeted to NMD as expected, but rather degraded by the nuclear turnover system. While most *S. cerevisiae* intron containing genes have only one intron, there are a few numbers of genes that contain two small introns; amongst these are the genes *MAT*a1, *DYN2*, *SUS1*, and *YOS1*. The degradation of the exon2 skipped forms of these transcripts was found to be dependent on the nuclear RNA turnover pathway consisting of the 5' to 3' exonuclease Rat1p and the nuclear exosome. *MAT*a1 was additionally found to be cleaved by the nuclear RNase III endonuclease Rnt1p. These findings show that nuclear degradation mechanisms have also evolved to complement the role of NMD to limit the accumulation of mRNAs that are erroneously spliced by the splicing machinery.

The dissertation of Tadashi Ryan Kawashima is approved.

Douglas L. Black

James W. Gober

Guillaume F. Chanfreau, Committee Chair

University of California, Los Angeles

2013

TABLE OF CONTENTS   **PAGE**

LIST OF FIGURES **PAGE**

ACKNOWLEDGMENTS

VITA

2000   University of California, at San Diego
B.S., Department of Chemistry and Biochemistry

2000   The Scripps Research Institute, Department of Neurobiology
Research Technician III

2002   GeneOhm Sciences Inc.
Research Associate III/ Interim Project Leader

2004   BD GeneOhm
Scientist

2007   University of California, at Los Angeles
Biochemistry and Molecular Biology Graduate Program

CHAPTER 1

Introduction to the

Nonsense-Mediate mRNA Decay Pathway

in *Saccharomyces cerevisiae*

First discovered in *Saccharomyces cerevisiae* over three decades ago (Losson and Lacroute 1979), the nonsense-mediated mRNA decay (NMD) pathway is a conserved eukaryotic quality control mechanism that rapidly degrades abnormal RNAs harboring premature translation termination codons (PTCs) (Chang, Imam et al. 2007). These PTCs have been shown to result from nonsense or frameshift mutations within the transcript. Degradation of such transcripts prevents the translation and accumulation of truncated proteins with potentially deleterious dominant negative, gain-of-function, or loss of function effects (Baserga and Benz 1992; Culbertson 1999). In fact, it is believed that approximately one third of known inherited diseases and many forms of cancers are caused by PTC generating nonsense or frameshift mutations (Dietz, McIntosh et al. 1993; Hall and Thein 1994; Dietz, Francke et al. 1995; Culbertson 1999; Frischmeyer and Dietz 1999; Holbrook, Neu-Yilik et al. 2004; Karam, Wengrod et al. 2013). Based on the wide range of diseases associated with PTCs, it is no surprise to speculate that furthering our understanding of the NMD mechanism may lead to the development of rational strategies for treatment.

UPF Proteins: Key NMD *Trans*-Acting Factors:

In *S. cerevisiae*, the core NMD machinery consist of three trans-acting factors, Upf1p, Upf2p, and Upf3p which destabilize nonsense containing transcripts (PTC+) without affecting the decay rate of most wildtype mRNAs (Muhlrad and Parker 1994; Cui, Hagan et al. 1995; He and Jacobson 1995; Lee and Culbertson 1995; Caponigro and Parker 1996; Gonzalez, Bhattacharya et al. 2001; Maquat and Serin 2001; Chang, Imam et al. 2007). Inactivation of any one of these factors results in the inhibition of NMD and the stabilization of PTC+ transcripts (Maderazo, He et al. 2000). High-density tiling microarrays have shown that global changes in the yeast transcriptome caused by loss of *UPF* gene function affected the accumulation of hundreds of mRNAs and the same mRNAs were affected regardless of which of the three *UPF* genes were inactivated (Lelivelt and Culbertson 1999; Amrani, Dong et al. 2006; Sayani, Janis

et al. 2008). These three core NMD components interact with one another to form a heterotrimeric complex known as the surveillance complex, which marks the transcript for rapid degradation (He, Brown et al. 1997; Shirley, Lelivelt et al. 1998).

UPF1 encodes a 971AA protein and is the most studied of the three Upf proteins. Studies have identified two major domains: 1) ATPase/helicase domain required for ATP-binding, ATP-independent nucleic acid-binding, nucleic acid-dependent ATP hydrolysis, and ATP-dependent helicase activity 2) N-terminal CH-domain required for Upf2p binding, ubiquitination function, and ribosome interaction mediated by Rps26p (Czaplinski, Weng et al. 1995; Weng, Czaplinski et al. 1996; Takahashi, Araki et al. 2008; Min, Roy et al. 2013). Upf1p's ATPase/helicase domain has a RNA binding motif that binds nucleic acids in the absence of ATP, but surprisingly, ATP binding by Upf1p results in the dissociation of the Upf1p:RNA complex, independent of actual ATP hydrolysis (Weng, Czaplinski et al. 1996). This ATP binding is dependent on a highly conserved lysine found within the ATPase/helicase domain of Upf1p as mutations of this lysine (K436) has shown to inhibit ATP binding as well as stabilize PTC+ mRNA (Weng, Czaplinski et al. 1996). Mechanistically, Upf1p is initially anticipated to bind tightly to its PTC+ RNA substrate. Once it interacts with Upf2p and Upf3p to form the surveillance complex, there is a conformational change in Upf1p which results in less extensive binding of RNA and allows the helicase domain to now be accessible for ATP binding and promote helicase activity to switch from RNA-clamping to RNA-unwinding mode (Chakrabarti, Jayachandran et al. 2011). The ATP dependent, Upf1p mediated unwinding of the PTC+ transcript results in the disassembly of stabilizing mRNPs and render the transcript susceptible to degradation (Franks, Singh et al. 2010).

Interestingly, initial experiments using mutations of conserved residues in these Upf1p regions sometimes resulted in the stabilization of PTC+ transcripts while supporting nonsense suppression (read-through of premature termination codon), while other mutations resulted in

both the inactivation of PTC+ degradation as well as prevent nonsense suppression (Weng, Czaplinski et al. 1996; de Pinto, Lippolis et al. 2004). These results implicated Upf1p in modulating both translation termination and mRNA turnover, and that these two functions are separable by mutations in certain regions of Upf1p (Weng, Czaplinski et al. 1996; Weng, Czaplinski et al. 1996). However, a more recent study suggests that nonsense suppression may be the result of an indirect effect of NMD inactivation. NMD inactivation was shown to stabilize the NMD sensitive *ALR1* mRNA (which encodes a plasma membrane Mg^{2+} transporter) and result in elevated cellular levels of Mg^{2+} (Johansson and Jacobson 2010). This high intracellular Mg^{2+} concentration has in the past been shown to affect translational read-through of a stop codon, although the exact mechanism is still unknown (Szer and Ochoa 1964; Capecchi 1967; Schlanger and Friedman 1973). One speculation is that the high concentration of Mg^{2+} causes a conformational distortion to the ribosome, which is itself an Mg^{2+}-dependent structure (Johansson and Jacobson 2010). Thus, the current view on Upf1p function is that it only affects mRNA turnover.

Yeast two-hybrid assays and co-immunoprecipitation experiments have also identified physical interactions between Upf1p's C-terminus (which overlap with the ATPase/helicase domain) and the two nuclear pore proteins (Nup) Nup100p and Nup116p (Nazarenus, Cedarberg et al. 2005). These two Nup proteins are homologous gene pairs that likely resulted from a gene duplication event and have been shown to localize to the cytoplasmic side of the nuclear pore to participate in mRNA transport (Wente and Blobel 1993; Strawn, Shen et al. 2001; Suntharalingam and Wente 2003). This interaction suggests that Upf1p is primed to associate with newly synthesized mRNA as it is transported from the nucleus to the cytoplasm in close temporal and spatial proximity to the association of the first ribosome (Culbertson and Neeno-Eckwall 2005; Nazarenus, Cedarberg et al. 2005).

Upf2p has been identified by two-hybrid screens for interaction partners of Upf1p via Upf1p's CH-domain (He and Jacobson 1995; He, Brown et al. 1996). Further analysis has identified the C-terminal region of Upf2p important for this Upf1p interaction (He and Jacobson 1995; He, Brown et al. 1996). This Upf1p:Upf2p interaction is required for rapid decay of PTC containing transcripts as point mutations which disrupted the binding of the two Upf proteins also inhibited NMD (He, Brown et al. 1996). In addition to this Upf1p interaction domain of Upf2p, an adjacent acidic region has been shown to affect NMD via interaction with Upf3p (He, Brown et al. 1996; Kadlec, Izaurralde et al. 2004). The N-terminal portion of Upf2p encodes a putative nuclear localization signal (NLS) which is also necessary for NMD as deletions of this region stabilized PTC+ transcripts (He, Brown et al. 1996). However, Upf2p has been reported to be cytoplasmically localized at steady state conditions and detailed experimental analysis on the functionality of the NLS has not been reported to date (Kadlec, Izaurralde et al. 2004; Culbertson and Neeno-Eckwall 2005). Thus, there is no conclusive evidence in support for the nuclear shuttling of Upf2p and further investigation is still necessary to support such claims.

In higher eukaryotes, it has been shown that both Upf1p and Upf2p undergo phosphorylation and dephosphorylation cycles and this ability is deemed crucial for NMD, as blockage of either event largely prevents NMD (Page, Carr et al. 1999; Denning, Jamieson et al. 2001; Pal, Ishigaki et al. 2001; Yamashita, Ohnishi et al. 2001; Anders, Grimson et al. 2003; Ohnishi, Yamashita et al. 2003; Wilkinson 2003; Grimson, O'Connor et al. 2004; Unterholzner and Izaurralde 2004; Johns, Grimson et al. 2007). In these higher eukaryotes, Smg1p has been shown to directly phosphorylate Upf1p while Smg5-7p proteins are known to promote dephosphorylation by protein phosphatase 2A (PP2A) (Yamashita, Ohnishi et al. 2001; Anders, Grimson et al. 2003; Chiu, Serin et al. 2003; Ohnishi, Yamashita et al. 2003). There is evidence to support the notion that in *S. cerevisiae*, Upf1p and Upf2p may also undergo phosphorylation (de Pinto, Lippolis et al. 2004; Wang, Cajigas et al. 2006). Western blot analysis identified a

slower migrating band in some *upf1* mutant alleles which were not detectable in the wildtype control unless in the presence of phosphatase inhibitors (de Pinto, Lippolis et al. 2004). The subsequent addition of excess lambda protein phosphatase once again converted the slow migrating Upf1p into the faster migrating isoform, thus confirming Upf1p can indeed be phosphorylated (de Pinto, Lippolis et al. 2004). The stable phosphorylated state initially observed for *upf1* mutant alleles was suggested to be the result of a slower rate of dephosphorylation than in the wildtype (de Pinto, Lippolis et al. 2004). Additional support for Upf1p phosphorylation comes from the identification of yeast gene *EBS1*, which has similar structure to the human dephosphorylation factor Smg7p (Luke, Azzalin et al. 2007). Yeast Ebs1p has been shown to interact with the helicase domain of Upf1p and its deletion has been shown to stabilize NMD targets (Chiu, Serin et al. 2003). Although Ebs1p appears to be required for NMD, one caveat is that no experiments have actually attempted to show if it actually affects the phosphorylation state of Upf1p. Thus, further studies are still required to determine whether Upf1p undergoes phosphorylation/dephosphorylation cycles as part of the NMD pathway, as reported for higher eukaryotes (Page, Carr et al. 1999; Pal, Ishigaki et al. 2001). Upf2p's ability to be phosphorylated was demonstrated by growing wildtype and *upf2Δ* strains in low-phosphate medium followed by [^{32}P] orthophosphate labeling (Wang, Cajigas et al. 2006). When these cytoplasmic extracts were immunoprecipitated with anti-Upf2 and ran on SDS-PAGE, wildtype extracts displayed a band not detectable in *upf2Δ*, indicating that Upf2p was phosphorylated (Wang, Cajigas et al. 2006). Additionally, incubation with calf intestine phosphatase resulted in Upf2p dephosphorylation and a faster migrating band in western blot experiments (Wang, Cajigas et al. 2006). Furthermore, site directed mutagenesis identified key serine residues in the N-terminal region of Upf2p that are required for phosphorylation as well as NMD activity (Wang, Cajigas et al. 2006). Thus, although still in the early stages of investigation, it appears that both Upf1p and Upf2p can be phosphorylated.

Upf3p contains both nuclear localization signals (NLS) as well as a nuclear export signal (NES) (Shirley, Lelivelt et al. 1998; Shirley, Ford et al. 2002; Shirley, Richards et al. 2004). Cytoplasmic reporters localized to the nucleus when fused to any of the three NLS of Upf3p and a nuclear reporter localized to the cytoplasm when fused to the NES of Upf3p (Shirley, Lelivelt et al. 1998). Immunofluorescence microscopy studies with Upf3p have shown that the nuclear export signal is very efficient, resulting in cytoplasmic staining (Shirley, Lelivelt et al. 1998). However, when Upf3p is over expressed, or when Upf3p's NES is mutated, Upf3p can be visualized in the nucleus, confirming shuttling across the nuclear envelope (Shirley, Lelivelt et al. 1998; Shirley, Ford et al. 2002). This nuclear shuttling of Upf3p is dependent on Srp1p (importin), as two-hybrid screen has identified it as having Upf3p interactions via Upf3p's NLS motifs, and overexpressed Upf3p was cytoplasmically localized in a temperature sensitive mutant of Srp1p (Shirley, Ford et al. 2002). Nuclear trapping of Upf3p caused by NES mutations resulted in NMD inhibition, indicating the importance of cytoplasmic export of Upf3p for NMD function and possibly suggests that mRNAs recruit Upf3p prior to or during their export to the cytoplasm (Shirley, Lelivelt et al. 1998; Culbertson 1999; Shirley, Ford et al. 2002; Shirley, Richards et al. 2004). The importance of Upf3p:Upf2p interaction for NMD was highlighted when overexpression of Upf2p was able to partially restored NMD function of *upf3* point mutants while such rescues were not observed in *upf3Δ* mutants (Shirley, Ford et al. 2002; Shirley, Richards et al. 2004). Two-hybrid confirmed that the point mutations of Upf3p resulted in weaker interactions with Upf2p and thus higher concentrations of Upf2p was required to have a large enough population of Upf2p:Upf3p for proper NMD function (Shirley, Ford et al. 2002; Shirley, Richards et al. 2004).

The cellular levels of Upf proteins were determined to elucidate whether a stoichiometric surveillance complex is able to form, or if the complex is more of a transient event. By using western blot of crude cellular extracts and purified proteins as standards, Upf1p was found to be

the most abundant of the three factors with approximately 1600 Upf1p/cell (Maderazo, He et al. 2000). This was followed by Upf2p at 1/10th the amount at 160 Upf2p/cell, and Upf3p was the least abundant at 80 Upf3p/cell (Maderazo, He et al. 2000). In contrast, global expression analysis of a fusion protein library showed approximately 6100 Upf1p/cell while levels of Upf3p and Upf2p were found at similar levels of approximately 1250/cell (Ghaemmaghami, Huh et al. 2003). Regardless of the method used to quantitate *UPF* protein levels, the amounts are found at concentrations much lower than the cellular levels of ribosomes (200,000/cell), release factors, or Xrn1p (Hereford and Rosbash 1977; Ghaemmaghami, Huh et al. 2003; von der Haar 2008). The reported levels of Upf proteins exclude the possibility of formation of a stable complex of Upf factors that associate with all ribosomes but rather supports a model where their interaction with the ribosome is more transient, possibly limited to ribosomes recognizing termination codons or during the pioneer round of translation (Maderazo, He et al. 2000).

Translation Requirement for NMD:

As implied up until now, NMD requires a translating ribosome to encounter a PTC. There is strong evidence to support this requirement of translation for NMD activation. These findings include: 1) NMD is inhibited by drugs (cycloheximide) and mutations that block translation initiation (Herrick, Parker et al. 1990; Peltz, Donahue et al. 1992; Zhang and Schwer 1997; Zuk and Jacobson 1998; Welch and Jacobson 1999), 2) nonsense-containing mRNAs are polysome associated, with size reflecting the position of the PTC in context to the open reading frame (He, Peltz et al. 1993; Maderazo, Belk et al. 2003; Hu, Petzold et al. 2010), 3) NMD can be partially prevented by nonsense-suppressing tRNAs (Losson and Lacroute 1979; Gozalbo and Hohmann 1990), 4) rapid turnover of nonsense-containing mRNA resumes immediately after removal of cycloheximide from media (Zhang, Welch et al. 1997), 5) NMD is functional only after at least one round of translation initiation/termination cycle has been completed (Ruiz-Echevarria, Gonzalez et al. 1998; Ruiz-Echevarria, Yasenchak et al. 1998), 6) Factors essential

for NMD interact with translation termination factors Sup45p/eRF1 and Sup35p/eRF3 (Czaplinski, Ruiz-Echevarria et al. 1998; Wang, Czaplinski et al. 2001), and 7) *ASH1* PTC+ mRNA is transported to the yeast bud tip by actin cables in a translationally repressed state, NMD only occurs when the transcript is relieved of its translational repression upon reaching the bud tip (Zheng, Finkel et al. 2008). This translational dependence of NMD pinpoints it as a cytoplasmic process where functional ribosomes reside. NMD as a cytoplasmic process is also supported by the cytoplasmic localization of the UPF proteins, as well as PTC+ reporter mRNAs being visualized in the cytoplasm by fluorescent in situ hybridization, and PTC+ transcripts defective in cytoplasmic export being reported as NMD insensitive (Kuperwasser, Brogna et al. 2004; Sheth and Parker 2006; Kurosaki and Maquat 2013).

From the extensive list of evidence above, it is clear that translation is required for a PTC+ transcript to be susceptible to NMD. In higher eukaryotes, it has been suggested that NMD is functional only during the pioneer round of translation, after which, a PTC+ transcript that escaped initial detection, is now stabilized (Maquat, Tarn et al. 2010). This, however, is not the case with *Saccharomyces cerevisiae* (Maderazo, Belk et al. 2003; Keeling, Lanier et al. 2004; Plant, Wang et al. 2004; Gao, Das et al. 2005; Amrani, Dong et al. 2006). Experimental evidence was provided by following the decay rates of PTC+ transcripts when the *UPF* genes are under the control of a galactose-inducible promoter (Maderazo, Belk et al. 2003). If the decay of the PTC+ transcript is tied to the pioneer round of translation, one would expect that the population of PTC+ mRNA transcribed in the absence of Upf1p should still be stabilized even after the induction of *UPF1*. However polysome profiling and northern blot results indicated that this was not the case, polysome association was observed for PTC+ prior to *UPF1* gene induction, but after induction, low polysome association was observed which indicated that the entire population of PTC+ transcripts was degraded rapidly upon *UPF1* induction (Maderazo, Belk et al. 2003). Further support for NMD not being limited to the pioneer

round of translation came from experiments following the decay profile of -1 programmed ribosome frame shift (PRF) targets. -1PRF with a low level of frameshift (~2-6% of transcript population) was observed to have strong effects on mRNA stability; if NMD was restricted to the pioneer round of translation, a larger population of -1PRF mRNA should have been stabilized (Plant, Wang et al. 2004). Computational modeling and decay profile analysis confirmed that NMD remained active beyond the pioneer round of translation by displayed a biphasic mode of decay (Plant, Wang et al. 2004). Another group used the cap status and decay rates of PTC+ transcripts to demonstrate that NMD is not limited to the pioneer round of translation. Newly transcribed mRNA is first bound by the primarily nuclear, but shuttling, cap-binding complex subunit, Cbc1p. Upon cytoplasmic transport of the transcript and after the pioneer round of translation, this cap-binding complex is replaced by the primarily cytoplasmic eIF-4E (Keeling, Lanier et al. 2004; Gao, Das et al. 2005). Immunoprecipitation of cellular extracts with Anti-Cbc1p and Anti-eIF-4E antibodies followed by RT-PCR for PTC+ transcripts indicated that PTC+ transcripts contain both capped states (Gao, Das et al. 2005). Additionally, strains lacking Cbc1p still maintained active NMD and showed the rapid decay of eIF4E-bound PTC+ transcripts (Kuperwasser, Brogna et al. 2004; Gao, Das et al. 2005). These observations suggest that PTC+ mRNAs that escape NMD during the pioneer round of translation have their Cbc1p cap replaced with eIF-4E, and that these transcripts can still be recognized as PTC+ and degraded by NMD in subsequent rounds of translation (Keeling, Lanier et al. 2004; Kuperwasser, Brogna et al. 2004).

Surveillance Complex Formation:

Once the ribosomes encounter of a premature translation termination codon and terminate translation, the surveillance complex composed of the three key NMD factors Upf1p, Upf2p, and Upf3p forms on the transcript to mark it for degradation. The spatial and temporal process of factor recruitment and surveillance complex formation is not fully understood at this

time, but a logical working model is as follows. Due to Upf3p's ability to shuttle between the nucleus and the cytoplasm, and the limited molecules of Upf3p molecules per cell, Shirley et al. puts forth the idea that Upf3p may associate to newly transcribed RNA in the nucleus by binding other mRNP proteins, since Upf3p lacks any RNA binding abilities on its own (Shirley, Ford et al. 2002; Kadlec, Izaurralde et al. 2004). This Cbc1p bound transcript is then exported to the cytoplasm to undergo the pioneer round of translation (Keeling, Lanier et al. 2004). During this process, the first translating ribosome displaces Upf3p along with other mRNPs on normal, PTC-free transcripts. In the case of the ribosome halting translation prematurely, the Upf3p remains bound to the transcript and seeds the formation of an Upf3p:Upf2p complex (He, Brown et al. 1996; Kadlec, Izaurralde et al. 2004). Upf1p is also recruited to the paused ribosome through interactions with termination release factors Sup45/eRF1 and Sup35/eRF3, which are recruited to every terminating ribosome regardless of PTC presence (Czaplinski, Ruiz-Echevarria et al. 1998; Czaplinski, Ruiz-Echevarria et al. 1999). The presence of all three Upf proteins results in the formation of the surveillance complex and activates NMD.

This working model also suggests that translation outside the pioneer round, in which the mRNPs have been displaced with the translating ribosome, the NMD factors must reassemble in subsequent rounds of translation (Culbertson and Neeno-Eckwall 2005). Upon PTC recognition by the ribosome, Sup45/eRF1 and Sup35/eRF3 are recruited to the stalled ribosome either pre-bound to Upf1p or recruiting Upf1p after the fact. After Upf1p:Sup45p:Sup35p subcomplex formation, Upf2p:Upf3p heteroduplex associates via Upf1p:Upf2p interaction and/or with Upf2p:Upf3:Sup35p/eRF3 interactions to form the surveillance complex (Gonzalez, Bhattacharya et al. 2001; Wang, Czaplinski et al. 2001).

Faux 3' UTR Model for PTC Identification:

All stop codons are initially recognized by the canonical translational termination factors Sup45p/eRF1 and Sup35p/eRF3. What then distinguishes a 'premature' stop codon from an

authentic one may be the presence of an extended 3'-UTR. In support of this, wildtype transcripts with long 3'-UTRs have also been observed to be marked for NMD degradation (Kebaara and Atkin 2009; Kuroha, Tatematsu et al. 2009). The typical range of 3'-UTR lengths in *S. cerevisiae* is from 50 to 200nt, with a median length of 121nt; thus 350nt was arbitrarily used as the cutoff to define a target as having a long 3'-UTR (Graber, Cantor et al. 1999; Kebaara and Atkin 2009). A global approach to identify wildtype transcripts with long 3'-UTRs identified 56 transcripts, and 91% of those that were tested (11 examples) were found to degraded by NMD (Kebaara and Atkin 2009). A long 3'-UTR was found to be sufficient to target a wildtype transcript for degradation as shown by replacing the 3'-UTR of a NMD–insensitive mRNA with the 750nt long 3'-UTR of *PGA1* mRNA (Kebaara and Atkin 2009). Supporting evidence is also provided from observations showing that 5' proximal PTCs, which mimics a long 3'-UTR, promote decay more readily than those near the 3' end (Losson and Lacroute 1979; Leeds, Peltz et al. 1991; Peltz, Brown et al. 1993; Hagan, Ruiz-Echevarria et al. 1995; Jacobson and Peltz 1996; Muhlrad and Parker 1999; Meyer, Plass et al. 2011). However, one caveat must be made in regards to the 3'-UTR length requirement to elicit NMD; one should consider the "effective" and not actual 3'-UTR length. RNA secondary structure in the long 3'-UTR has been shown to shorten their "effective" length and thus suppress NMD (Eberle, Stalder et al. 2008).

These extended 3'-UTR experiments provided the beginnings for the *faux* 3'-UTR model for NMD put forth by Amrani et al.; where a normal transcript is identified through interactions between the ribosome and mRNPs localized 3' of the stop codon (Amrani, Ganesan et al. 2004). Her experiments showed that translation termination at a PTC is aberrant as toeprint assay for ribosome stalling was identifiable in PTC+ mRNA while no toeprints were observed under normal termination; termination at a PTC is thus much less efficiently released (Amrani, Ganesan et al. 2004; Amrani, Dong et al. 2006). These aberrant toeprints are linked to NMD as

ufp1Δ strains fail to accumulate toeprints at PTCs (Amrani, Ganesan et al. 2004; Amrani, Dong et al. 2006). This stalling was also shown to be dependent of 3'-UTR length as the presence of ribosome stalling was no longer observed when PTC+ transcript was fused to a shorter, normal 3' UTR (Amrani, Ganesan et al. 2004; Amrani, Dong et al. 2006). This effect was mimicked by a long 3'-UTR transcript tethered with Pab1p (Amrani, Dong et al. 2006). Additionally, tethering of Sup35p/eRF3, a translation termination factor known to physically interact with Pab1p, was also able to stabilize PTC+ mRNA (Amrani, Dong et al. 2006). These results led Amrani et al. to suggest that interaction between the terminating ribosome, Sup35p/eRF3, and 3' UTR mRNP Pab1p are important in marking a transcript as a normal transcript. The *faux* 3'-UTR model proposes that when termination occurs at a PTC too far away from this 3'-UTR mRNP environment, disassembly of the ribosome is slow, presumably because in this spatial environment, they cannot receive the proper termination stimulating signal (Amrani, Ganesan et al. 2004). As a result, Upf1p is recruited to the terminating ribosome through interaction with Sup45p/eRF1 and Sup35p/eRF3 (Amrani, Ganesan et al. 2004). The competition between Pab1p and Upf1p for interaction with Sup35p/eRF3 enables the cell to distinguish between correct and aberrant translation termination (Eberle, Stalder et al. 2008; Kervestin, Li et al. 2012).

To lend additional support for the important role of Pab1p and to determine if mRNPs other than Pab1p that are deposited during 3' end formation are needed, a synthetic PTC+ transcript which forms it's poly(A) tail without the requirement of normal 3' end forming machinery, by introducing a self-cleaving ribozyme sequence downstream of a poly(A) sequence, was tested (Baker and Parker 2006). This transcript was found to be targeted for NMD degradation even if 3' end forming mRNPs are not associated with it. This construct however, was still found to bind Pab1p, and thus argue that Pab1p alone is sufficient and a

mRNP complex deposited during 3' end formation may be dispensable (Baker and Parker 2006).

The *faux* 3'-UTR model hinges upon the interaction between Sup35p/eRF3 with Upf1p and Pab1p as a major determinant in distinguishing a normal versus abnormal target. Competition between Upf1p and Pab1p for binding Sup35p/eRF3 was tested with purified recombinant proteins, where sequential addition of Pab1p and Upf1p with Sup35p/eRF3 showed limited support for direct competition between Upf1p and Pab1p (Kervestin, Li et al. 2012). Binding of Upf1p was not affected by preincubation of Sup35p/eRF3p with Pab1p even though Pab1p was shown to associate with Sup35p/eRF3 at 5x lower concentrations than Upf1p (Kervestin, Li et al. 2012). Additionally, the displacement of bound Pab1p was not observed; indicate that binding of Pab1p and Upf1p to Sup35p/eRF3 are not mutually exclusive (Kervestin, Li et al. 2012). However, Pab1p binding was shown to be inhibited by preincubation of Sup35p/eRF3p and Upf1p (Kervestin, Li et al. 2012). Thus, suggesting that Pab1p and Upf1p do not bind two entirely different sites on Sup35p/eRF3p and that the two binding sites may be overlapping. In contrast to sequential addition of factors, assays monitoring simultaneous additions showed interactions more consistent with competition of Pab1p and Upf1p for Sup35p/eRF3p binding. Simultaneous presence of Upf1p was able to block Pab1p binding to Sup35p/eRF3 while high concentrations of Pab1p affected Upf1p:Sup35p/eRF3p interactions; cellular levels of Pab1p is very high and consistent with results (von der Haar 2008; Kervestin, Li et al. 2012). Based on these experimental results, the competing interactions between Upf1p and Pab1p for Sup35p/eRF3 does not appear to be as simple as originally hypothesized. Further experiments are still necessary to elucidate the complexities and resolve some issues in regards to the competing interactions amongst these factors.

Modes of Degradation:

Decay of Normal Transcripts:

Normal turnover of mRNAs is deadenylation-dependent and requires multiple rounds of deadenylation prior to 5'-decapping and subsequent degradation by the cytoplasmic 5' to 3' exonuclease Xrn1p (Decker, Parker et al. 1994; Beelman, Stevens et al. 1996; Caponigro and Parker 1996). The poly(A) tail is initially degraded until it is too short for the binding of poly(A)-binding protein (Pab1p) (Caponigro and Parker 1995). At this point, the decapping enzyme complex composed of Dcp1p-Dcp2p is recruited to the 5' cap structure of the mRNA and removes the 7-methylguanylate cap from the mRNA (Beelman, Stevens et al. 1996). The now unprotected transcript is accessible by the cytoplasmic 5' to 3' exonuclease Xrn1p for rapid digestion (Decker, Parker et al. 1994; Caponigro and Parker 1996).

Major Deadenylation Independent 5' to 3' Decay of PTC Transcripts:

In contrast to normal mRNA turnover, the presence of a PTC bypasses the rate-limiting step of deadenylation and result in rapid degradation of aberrant transcripts (Muhlrad and Parker 1994; Caponigro and Parker 1996). In yeast, in addition to the three Upf proteins, rapid degradation of PTC+ transcripts require the decapping enzymes Dcp1p and Dcp2p as well as the 5' to 3' exonuclease Xrn1p (Muhlrad and Parker 1994; Dunckley and Parker 1999; Maderazo, He et al. 2000; He and Jacobson 2001; He, Li et al. 2003; Hu, Petzold et al. 2010). *UPF* deletion mutants have been shown to exhibit accumulation of capped, full length PTC+ transcripts while cells lacking Xrn1p showed accumulation of mRNA lacking a cap structure (He and Jacobson 2001). When *XRN1* was deleted in *upf* mutants, the double mutants showed that the PTC+ mRNA was capped and full length, similar to *upf* single mutants (He and Jacobson 2001). These results support a mechanism of degradation where Upf1p mediates decapping of a PTC+ mRNA followed by degradation by the 5' to 3' exonuclease Xrn1p (He and Jacobson 2001). This connection between the *UPF* genes, Dcp1p, and Xrn1p was observed in high

density microarrays experiments. Global transcript levels showed *UPF* gene deletions had huge overlap with each other as well as with deletions of *XRN1* or *DCP1* (He, Li et al. 2003). Additional evidence in support of this mechanism of Upf protein directed decapping comes from two-hybrid results where Upf1p was found to interact with enhancer of decapping protein, Edc3p, which mediates binding of the decapping enzyme Dcp2p (Swisher and Parker 2011). However, it was shown that Edc3p was not essential for NMD function; deletion of *EDC3* still showed PTC+ reporter decay rates similar to wildtype (Swisher and Parker 2011). This suggests that the Edc3p:Upf1p interaction is not required for the recruitment of Dcp2p and possible explanations for this include 1) Edc3p:Upf1p interaction is necessary for only a subset of NMD transcripts, 2) Edc3p may be essential for promoting NMD when environmental or growth conditions are altered, 3) there are additional factors which may recruit Dcp2p to the NMD substrate in the absence of Edc3p, 4) Upf1p and Edc3p may also be involved in another decay pathway unrelated to NMD (Swisher and Parker 2011). Further experiments are still required to resolve exactly how Upf proteins direct the decapping of PTC+ transcripts.

This degradation of PTC+ transcripts has been suggested to occur in cytoplasmic foci knows as processing bodies (P-bodies) and Upf1p is sufficient for P-body targeting while Upf2p and Upf3p act downstream of P-body targeting to trigger decapping (Sheth and Parker 2006). GPF tagged Upf proteins were all uniformly distributed in the cytoplasm until enzymes involved in decay were deleted (*dcp1∆, dcp2∆, and xrn1∆*), in which case localization to cytoplasmic foci known as P-bodies was observed (Sheth and Parker 2006). This indicates that under normal steady state conditions, Upf proteins are rapidly released from rapidly degrading PTC+ mRNA and do not accumulate to high levels. The degradation of transcripts at P-bodies was shown to be specific to PTC+ transcripts as: 1) deletion of factors associated with decapping of only normal transcripts did not result in P-body formation 2) PTC+ transcripts bound to GFP showed P-body localization while PTC- control transcripts did not (Sheth and Parker 2006). This PTC+

transcript localization to P-bodies was dependent on the presence of Upf1p as PTC+ transcripts, which accumulated at much higher levels in the absence of Upf1p, showed decreased levels of P-bodies in the *upf1Δ* strain (Sheth and Parker 2006). Upf2p and Upf3p deletion mutants were still able to target PTC+ transcripts to P-bodies suggesting that while necessary for NMD, they are not required for P-body targeting, and thus act downstream in the decay process (Sheth and Parker 2006).

Minor deadenylation dependent 3' to 5' Decay of PTC Transcripts:

In absence of Dcp1p, it has been shown that there is also an alternative decapping independent NMD pathway in which rapid deadenylation is followed by 3' to 5' degradation of PTC+ mRNA (He, Li et al. 2003; Mitchell and Tollervey 2003; Takahashi, Araki et al. 2003). This mode of decay holds true to NMD as it requires the presence of Upf proteins as well as translation (Mitchell and Tollervey 2003; Takahashi, Araki et al. 2003). PTC+ transcripts were shown to have accelerated rate of deadenylation followed by cytoplasmic exosome (complex of 3' to 5' exonucleases and associated Ski proteins) mediated decay (Mitchell, Petfalski et al. 1997; Anderson and Parker 1998; Allmang, Petfalski et al. 1999; Brown, Bai et al. 2000; Araki, Takahashi et al. 2001; Mitchell and Tollervey 2003; Takahashi, Araki et al. 2003). Recruitment of the exosome complex may be Upf1p mediated as it has been shown to physically interact with the N-terminal domain of Ski7p, a key component of the cytoplasmic exosome, suggesting that Ski7p acts as the bridge between NMD and 3' to 5' decay machinery (Takahashi, Araki et al. 2003).

Degradation of Truncated Protein Resulting from Nonsense Mutations:

NMD generally reduces the abundance of mRNAs to approximately 5 to 25% of normal, however, proteins from a reporter PTC+ transcript are found at much lower levels, down to even 1% (Kuroha, Tatematsu et al. 2009). This additional down regulation of protein appears to be

dependent on Upf1p as *UPF* knockout strains showed greater protein stability by measured half-life (Kuroha, Tatematsu et al. 2009). Proteasome mutant *rpn6-2* demonstrated that protein stability was the result of reduced proteasome targeting in *upf1Δ* (Kuroha, Tatematsu et al. 2009). Additionally, replacing the long NMD-sensitive 3'-UTR in the reporter with a shorter NMD-insensitive 3'-UTR resulted in protein stabilization, to levels similar to NMD inactivation via *upf1Δ* (Kuroha, Tatematsu et al. 2009). This all ties in with Upf1p's reported E3 ubiquitin ligase activity (Takahashi, Araki et al. 2008). Structural analysis of Upf1p has shown that the N-terminal CH domain forms a feature reminiscent of a RING-related E3 ubiquitin ligase (Kadlec, Guilligay et al. 2006; Takahashi, Araki et al. 2008). In support for ubiquitin ligase function, Upf1p's CH-domain was observed to interact directly with ubiquitin-conjugating enzyme Ubc3p (E2 protein) and in *in vitro* self-ubiquitination assays, wildtype Upf1p self-ubiquitinated and form polyubiquitin chains in the presence of added Ubc3p (Takahashi, Araki et al. 2008). Mutations of key cysteine and histidine residues of the CH-domain blocked ubiquitination and NMD function, hinting to a possible link between ubiquitination and NMD (Takahashi, Araki et al. 2008). One can imagine Upf1p recruitment to an aberrant transcript, followed by Upf1p:Ubc3p interaction to mediated Ubc3p dependent ubiquitination of the PTC+ translated protein. This ubiquitination marks the truncated protein for proteasome degradation to limit the potential adverse effects of having a dominant-negative or gain of function protein.

Difference in Decay Mechanisms between *S. cerevisiae* and Higher Eukaryotes:

In higher eukaryotes, the regulation of NMD requires SMG proteins in addition to the UPF proteins. These SMG proteins are involved in the phosphorylation cycle of UPF1 (Ohnishi, Yamashita et al. 2003; Grimson, O'Connor et al. 2004; Yamashita, Izumi et al. 2009). The cycling of phosphorylation states results in the recruitment of different mRNA decay enzymes which can promote two separate modes of decay (Ohnishi, Yamashita et al. 2003). The SMG6 dependent mode of decay uses its endonuclease activity to cleave PTC+ transcripts into two

halves (Huntzinger, Kashima et al. 2008; Eberle, Lykke-Andersen et al. 2009; Franks, Singh et al. 2010). The resulting 5' and 3' fragments are then degraded by the cytoplasmic exosome and Xrn1p, respectively (Gatfield and Izaurralde 2004; Huntzinger, Kashima et al. 2008). The second mode of decay recruits SMG5-7 and forgoes the endonucleolytic cleavage step (Fukuhara, Ebert et al. 2005). SMG5-7 recruitment promotes 5' to 3' and 3' to 5' decay of the PTC+ transcript, mediated by Dcp1p/Dcp2p, Xrn1p and the cytoplasmic exosome (Fukuhara, Ebert et al. 2005). The selection in the mode of decay for higher eukaryotes depends on its genetic makeup. *D. melanogaster* for example, lacks SMG7 and thus PTC+ transcripts are degraded in a SMG6 dependent manner (Gatfield, Unterholzner et al. 2003). In contrast, *C. elegans* and humans have SMG5-7, thus favoring the non-endonucleolytic mode of decay (Gatfield, Unterholzner et al. 2003).

NMD assures a high level of accuracy:

Premature translation termination codons can be generated by many different means such as 1) presence of upstream open reading frame (uORF), 2) leaky scanning of the ribosome, 3) presence of cis-acting programmed ribosome frame shift sequences (PRFs), and 4) translation of intron retained pre-mRNAs (Vilela, Linz et al. 1998; Vilela, Ramirez et al. 1999; Welch and Jacobson 1999; Ruiz-Echevarria and Peltz 2000; Kebaara, Nazarenus et al. 2003; Gaba, Jacobson et al. 2005). Regardless of their origin, these PTC+ transcripts are identified and rapidly degraded by NMD to assure a high level of accuracy necessary for proper gene expression. Additionally, it appears that some PTC encounters by the ribosome may be used as a means to rapidly down regulate gene expression under certain cellular conditions.

Translation initiation from an upstream open reading frame (uORF) can result in the ribosome stalling prior to the ORF, resulting in a long *faux* 3' UTR, marking the transcript for rapid degradation (Cui, Hagan et al. 1995; Vilela, Linz et al. 1998; Vilela, Ramirez et al. 1999; Ruiz-Echevarria and Peltz 2000; Messenguy, Vierendeels et al. 2002). Some known uORF

targets include *ALR1*, *CPA1*, *PPR1*, *GCN4*, and *YAP1* (Vilela, Linz et al. 1998; Vilela, Ramirez et al. 1999; Ruiz-Echevarria and Peltz 2000; Kebaara, Nazarenus et al. 2003; Gaba, Jacobson et al. 2005). In the case of *CPA1,* the use of its uORF is strongly enhanced in the presence of arginine in the growth media (Messenguy, Vierendeels et al. 2002; Gaba, Jacobson et al. 2005). The use of the uORF results in the translation of a short 25 amino acid polypeptide, which stalls the ribosome prior to the *CPA1* coding sequence, thus marking the transcript for NMD degradation (Werner, Feller et al. 1987; Delbecq, Werner et al. 1994; Messenguy, Vierendeels et al. 2002; Gaba, Jacobson et al. 2005). The gene product from *CPA1* encodes a protein subunit required for the arginine biosynthesis pathway and thus it is no surprise that the expression of *CPA1* ORF is inhibited in presence of sufficient arginine in the media. This example is a very interesting case where NMD is playing an important role in regulating gene expression.

PTCs can also arise from leaky scanning of the ribosome where usage of an internal AUG start codon results in a PTC generating frame-shift (Welch and Jacobson 1999; Johansson and Jacobson 2010). In the case of *SPT10*, the proper AUG is found in a poor translation initiation sequence context. This results in a fraction of ribosomes to scan past the first AUG to use a downstream +1 frame-shifted AUG within the ORF (Baim and Sherman 1988; Cigan, Pabich et al. 1988; Yun, Laz et al. 1996). Consistent with ribosome initiation at the downstream AUG site, introduction of another frameshift further downstream to shift the transcript back to the wildtype reading frame resulted in protein expression displaying partial function (Natsoulis, Dollard et al. 1991). Additionally, CCC mutation of the internal AUG sequence rendered the *SPT10* mutant transcript insensitive to NMD, suggesting that the AUG in the +1 reading frame was necessary to promote NMD (Welch and Jacobson 1999). Bioinformatics screen identified 18 additional candidates for leaky scanning; of these, *REV7*, *UBP7*, and *STE50* mRNA showed transcript stabilization in the *upf1Δ* NMD mutant (Welch and

Jacobson 1999). Based on these results, one can hypothesize that leaky scanning may provide yeast with a novel regulatory opportunity; any condition which increases the rate of leaky scanning can decrease transcript levels while conditions favoring initiation at the first AUG will result in mRNA stabilization.   Moreover, in higher eukaryotes, leaky scanning has been suggested as a mechanism to make multiple proteins from a single transcript, possibly differing in localization or function (Vilela, Ramirez et al. 1999).   It would be interesting if future experiments in *S. cerevisiae* followed up on these possibilities.

Programmed ribosome frameshift (PRF) are cis-acting sequences which redirect ribosomes into new reading frames and can induce a PTC (Plant, Wang et al. 2004; Belew, Advani et al. 2011). PRFs have been found to be employed by organisms representing every branch in the tree of life, suggesting that it's an ancient means to control the expression of actively translated mRNAs (Dinman 2006). Initial experiments used reporter constructs with a -1PRF to demonstrate that it does indeed act as a cis-acting destabilizing element for NMD (Plant, Wang et al. 2004). Employing -1PRFs of various efficiencies demonstrated that mRNA stability was inversely associated with the strength of the PRF (Plant, Wang et al. 2004). Additional supporting evidence surfaced when *EST2's* mRNA was shown to harbor a -1PRF that marked it for NMD degradation (Belew, Advani et al. 2011). In addition to -1 PRF, +1 programmed ribosome frameshift (+1PRF) of *EST3* mRNA was also determined to be a substrate for NMD (He, Li et al. 2003). Thus, both synthetic reporters and more importantly, *EST2* and *EST3* natural transcripts have shown that PRF elements which cause the ribosome to encounter a PTC are targeted for NMD degradation.

Intron retained pre-mRNAs have been shown to escape the nucleus and their subsequent translation results in the ribosome encountering an intronic PTC (He, Peltz et al. 1993; Li, Paulovich et al. 1995; Sayani, Janis et al. 2008). The global impacts of NMD inactivation on intron-containing transcripts have been shown to affect roughly 30% of all intron

containing genes by the use of high-density tiling arrays (Sayani, Janis et al. 2008). The importance of intron identity in NMD targeted decay was demonstrated by intron swapping experiments where an efficiently spliced gene, *GOT1*, had its intron swapped with inefficiently spliced *RPL19A* or *RPL11B* intronic sequence (Sayani, Janis et al. 2008). This swap resulted in the *GOT1* fusion to be inefficiently spliced and under the quality control of NMD, confirming that the intron is a key determinant (Sayani, Janis et al. 2008). Closer inspection of intronic features found that, in unspliced NMD targets, there is an enrichment of suboptimal splicing signals (Sayani, Janis et al. 2008). When these suboptimal splicing signals were mutated to the strong, consensus splice signal, efficient splicing was observed (Sayani, Janis et al. 2008). Thus, significant fraction of the intronome is affected by NMD inactivation and this is the result of transcripts containing suboptimal splicing signals.

Expanding the Role of NMD in *S. cerevisiae*:

What has been described above summarizes our current knowledge on the subject of nonsense-mediated mRNA decay in *Saccharomyces cerevisiae*. The following chapters cover my efforts in expanding our understanding of NMD in the quality control of splicing. The splicing of introns required the spliceosome to recognize the splice signals encoded within the intron and then correctly excise the intron and rejoin the exons through two transesterification reactions (Padgett, Grabowski et al. 1986; Sharp 1994; Du and Rosbash 2002; Lund and Kjems 2002; Jurica and Moore 2003; Freund, Hicks et al. 2005; Kent, Ritchie et al. 2005; Nilsen 2005; Sharp 2005). It has been shown above that unspliced pre-mRNAs that escape the nucleus are targeted for NMD degradation (He, Peltz et al. 1993; Sayani, Janis et al. 2008). Thus, it would be reasonable to speculate that mutations of splicing factors that are associated with the spliceosome may negatively affect the splicing of specific transcripts and thus increase the amount of unspliced pre-mRNAs. One can also envision that when the spliceosome inappropriately uses cryptic splice sites rather than the proper splice site, the resulting transcript

has a high probability of introducing a frame shift to generate a PTC+ transcript. In fact, there may be even be a regulatory component associated with alternative splice site selection, where the selection of an alternative splice site is favored under certain conditions as a means to rapidly decrease transcript levels by the NMD system. The exploration of these inquiries is discussed in detail in the following chapters.

References:

Allmang, C., E. Petfalski, et al. (1999). "The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases." <u>Genes Dev</u> **13**(16): 2148-2158.

Amrani, N., S. Dong, et al. (2006). "Aberrant termination triggers nonsense-mediated mRNA decay." <u>Biochem Soc Trans</u> **34**(Pt 1): 39-42.

Amrani, N., R. Ganesan, et al. (2004). "A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay." <u>Nature</u> **432**(7013): 112-118.

Anders, K. R., A. Grimson, et al. (2003). "SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A." <u>The EMBO journal</u> **22**(3): 641-650.

Anderson, J. S. and R. P. Parker (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." <u>The EMBO journal</u> **17**(5): 1497-1506.

Araki, Y., S. Takahashi, et al. (2001). "Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast." <u>The EMBO journal</u> **20**(17): 4684-4693.

Aronova, A., D. Bacíková, et al. (2007). "Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing." <u>RNA</u> **13**(9): 1437-1444.

Baim, S. B. and F. Sherman (1988). "mRNA structures influencing translation in the yeast Saccharomyces cerevisiae." <u>Mol Cell Biol</u> **8**(4): 1591-1601.

Baker, K. E. and R. Parker (2006). "Conventional 3' end formation is not required for NMD substrate recognition in Saccharomyces cerevisiae." <u>RNA</u> **12**(8): 1441-1445.

Baserga, S. J. and E. J. Benz, Jr. (1992). "Beta-globin nonsense mutation: deficient accumulation of mRNA occurs despite normal cytoplasmic stability." <u>Proc Natl Acad Sci U S A</u> **89**(7): 2935-2939.

Beelman, C. A., A. Stevens, et al. (1996). "An essential component of the decapping enzyme required for normal rates of mRNA turnover." <u>Nature</u> **382**(6592): 642-646.

Belew, A. T., V. M. Advani, et al. (2011). "Endogenous ribosomal frameshift signals operate as mRNA destabilizing elements through at least two molecular pathways in yeast." <u>Nucleic Acids Res</u> **39**(7): 2799-2808.

Bergkessel, M., G. B. Whitworth, et al. (2011). "Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast." <u>RNA</u> **17**(8): 1461-1478.

Brown, J. T., X. Bai, et al. (2000). "The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo." <u>RNA</u> **6**(3): 449-457.

Capecchi, M. R. (1967). "Polarity in vitro." <u>J Mol Biol</u> **30**(1): 213-217.

Caponigro, G. and R. Parker (1995). "Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast." <u>Genes Dev</u> **9**(19): 2421-2432.

Caponigro, G. and R. Parker (1996). "Mechanisms and control of mRNA turnover in Saccharomyces cerevisiae." <u>Microbiological reviews</u> **60**(1): 233-249.

Chakrabarti, S., U. Jayachandran, et al. (2011). "Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2." <u>Mol Cell</u> **41**(6): 693-703.

Chanfreau, G. F. (2010). "A dual role for RNA splicing signals." <u>EMBO Rep</u> **11**(10): 720-721.

Chang, Y. F., J. S. Imam, et al. (2007). "The nonsense-mediated decay RNA surveillance pathway." <u>Annual review of biochemistry</u> **76**: 51-74.

Chiu, S. Y., G. Serin, et al. (2003). "Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1." <u>RNA</u> **9**(1): 77-87.

Cigan, A. M., E. K. Pabich, et al. (1988). "Mutational analysis of the HIS4 translational initiator region in Saccharomyces cerevisiae." <u>Mol Cell Biol</u> **8**(7): 2964-2975.

Clifton, D. and D. G. Fraenkel (1981). "The gcr (glycolysis regulation) mutation of Saccharomyces cerevisiae." J Biol Chem **256**(24): 13074-13078.

Cui, Y., K. W. Hagan, et al. (1995). "Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon." Genes Dev **9**(4): 423-436.

Culbertson, M. R. (1999). "RNA surveillance. Unforeseen consequences for gene expression, inherited genetic disorders and cancer." Trends in genetics : TIG **15**(2): 74-80.

Culbertson, M. R. and E. Neeno-Eckwall (2005). "Transcript selection and the recruitment of mRNA decay factors for NMD in Saccharomyces cerevisiae." RNA **11**(9): 1333-1339.

Culbertson, M. R. and E. Neeno-Eckwall (2005). Transcript selection and the recruitment of mRNA decay factors for NMD in Saccharomyces cerevisiae. RNA. **11:** 1333-1339.

Czaplinski, K., M. J. Ruiz-Echevarria, et al. (1999). "Should we kill the messenger? The role of the surveillance complex in translation termination and mRNA turnover." Bioessays **21**(8): 685-696.

Czaplinski, K., M. J. Ruiz-Echevarria, et al. (1998). "The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs." Genes Dev **12**(11): 1665-1677.

Czaplinski, K., Y. Weng, et al. (1995). "Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation." RNA **1**(6): 610-623.

de Pinto, B., R. Lippolis, et al. (2004). "Overexpression of Upf1p compensates for mitochondrial splicing deficiency independently of its role in mRNA surveillance." Mol Microbiol **51**(4): 1129-1142.

Decker, C. J., R. Parker, et al. (1994). "Mechanisms of mRNA degradation in eukaryotes Mechanisms and control of mRNA turnover in Saccharomyces cerevisiae." Trends in biochemical sciences **19**(8): 336-340.

Delbecq, P., M. Werner, et al. (1994). "A segment of mRNA encoding the leader peptide of the CPA1 gene confers repression by arginine on a heterologous yeast gene transcript." Mol Cell Biol **14**(4): 2378-2390.

Denning, G., L. Jamieson, et al. (2001). "Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein." J Biol Chem **276**(25): 22709-22714.

Dietz, H., U. Francke, et al. (1995). "The question of heterogeneity in Marfan syndrome." Nat Genet **9**(3): 228-231.

Dietz, H. C., I. McIntosh, et al. (1993). "Four novel FBN1 mutations: significance for mutant transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome." Genomics **17**(2): 468-475.

Dinman, J. D. (2006). "Programmed Ribosomal Frameshifting Goes Beyond Viruses: Organisms from all three kingdoms use frameshifting to regulate gene expression, perhaps signaling a paradigm shift." Microbe Wash DC **1**(11): 521-527.

Du, H. and M. Rosbash (2002). "The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing." Nature **419**(6902): 86-90.

Dunckley, T. and R. Parker (1999). "The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif." The EMBO journal **18**(19): 5411-5422.

Eberle, A. B., S. Lykke-Andersen, et al. (2009). "SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells." Nat Struct Mol Biol **16**(1): 49-55.

Eberle, A. B., L. Stalder, et al. (2008). "Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region." PLoS Biol **6**(4): e92.

Franks, T. M., G. Singh, et al. (2010). "Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay." Cell **143**(6): 938-950.

Freund, M., M. J. Hicks, et al. (2005). "Extended base pair complementarity between U1 snRNA and the 5' splice site does not inhibit splicing in higher eukaryotes, but rather increases 5' splice site recognition." Nucleic Acids Res **33**(16): 5112-5119.

Frischmeyer, P. A. and H. C. Dietz (1999). "Nonsense-mediated mRNA decay in health and disease." Human molecular genetics **8**(10): 1893-1900.

Fukuhara, N., J. Ebert, et al. (2005). "SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway." Mol Cell **17**(4): 537-547.

Gaba, A., A. Jacobson, et al. (2005). "Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay." Mol Cell **20**(3): 449-460.

Gao, Q., B. Das, et al. (2005). "Cap-binding protein 1-mediated and eukaryotic translation initiation factor 4E-mediated pioneer rounds of translation in yeast." Proc Natl Acad Sci U S A **102**(12): 4258-4263.

Gasch, A. P., P. T. Spellman, et al. (2000). "Genomic expression programs in the response of yeast cells to environmental changes." Mol Biol Cell **11**(12): 4241-4257.

Gatfield, D. and E. Izaurralde (2004). "Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila." Nature **429**(6991): 575-578.

Gatfield, D., L. Unterholzner, et al. (2003). "Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways." The EMBO journal **22**(15): 3960-3970.

Ghaemmaghami, S., W. K. Huh, et al. (2003). "Global analysis of protein expression in yeast." Nature **425**(6959): 737-741.

Gonzalez, C. I., A. Bhattacharya, et al. (2001). "Nonsense-mediated mRNA decay in Saccharomyces cerevisiae." Gene **274**(1-2): 15-25.

Gottschalk, A., J. Tang, et al. (1998). "A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins." RNA **4**(4): 374-393.

Gozalbo, D. and S. Hohmann (1990). "Nonsense suppressors partially revert the decrease of the mRNA level of a nonsense mutant allele in yeast." Curr Genet **17**(1): 77-79.

Graber, J. H., C. R. Cantor, et al. (1999). "Genomic detection of new yeast pre-mRNA 3'-end-processing signals." Nucleic Acids Res **27**(3): 888-894.

Green, R. E., B. P. Lewis, 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.

Grimson, A., S. O'Connor, et al. (2004). "SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans." Mol Cell Biol **24**(17): 7483-7490.

Grund, S. E., T. Fischer, et al. (2008). "The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression." J Cell Biol **182**(5): 897-910.

Hagan, K. W., M. J. Ruiz-Echevarria, et al. (1995). "Characterization of cis-acting sequences and decay intermediates involved in nonsense-mediated mRNA turnover." Mol Cell Biol **15**(2): 809-823.

Hall, G. W. and S. Thein (1994). "Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia." Blood **83**(8): 2031-2037.

He, F., A. H. Brown, et al. (1996). "Interaction between Nmd2p and Upf1p is required for activity but not for dominant-negative inhibition of the nonsense-mediated mRNA decay pathway in yeast." RNA **2**(2): 153-170.

He, F., A. H. Brown, et al. (1997). "Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway." Mol Cell Biol **17**(3): 1580-1594.

He, F. and A. Jacobson (1995). "Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen." Genes Dev **9**(4): 437-454.

He, F. and A. Jacobson (2001). "Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs." Mol Cell Biol **21**(5): 1515-1530.

He, F., X. Li, et al. (2003). "Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast." Mol Cell **12**(6): 1439-1452.

He, F., S. W. Peltz, et al. (1993). "Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant." Proc Natl Acad Sci U S A **90**(15): 7034-7038.

Hereford, L. M. and M. Rosbash (1977). "Number and distribution of polyadenylated RNA sequences in yeast." Cell **10**(3): 453-462.

Herrick, D., R. Parker, et al. (1990). "Identification and comparison of stable and unstable mRNAs in Saccharomyces cerevisiae." Mol Cell Biol **10**(5): 2269-2284.

Holbrook, J. A., G. Neu-Yilik, et al. (2004). "Nonsense-mediated decay approaches the clinic." Nat Genet **36**(8): 801-808.

Holland, M. J., T. Yokoi, et al. (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**(2): 813-820.

Hu, W., C. Petzold, et al. (2010). "Nonsense-mediated mRNA decapping occurs on polyribosomes in Saccharomyces cerevisiae." Nat Struct Mol Biol **17**(2): 244-247.

Huntzinger, E., I. Kashima, et al. (2008). "SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan." RNA **14**(12): 2609-2617.

Isken, O. and L. E. Maquat (2007). "Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function." Genes Dev **21**(15): 1833-1856.

Jacobson, A. and S. W. Peltz (1996). "Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells." Annual review of biochemistry **65**: 693-739.

Jaillon, O., K. Bouhouche, et al. (2008). "Translational control of intron splicing in eukaryotes." Nature **451**(7176): 359-362.

Johansson, M. J. and A. Jacobson (2010). "Nonsense-mediated mRNA decay maintains translational fidelity by limiting magnesium uptake." Genes Dev **24**(14): 1491-1495.

Johns, L., A. Grimson, et al. (2007). "Caenorhabditis elegans SMG-2 selectively marks mRNAs containing premature translation termination codons." Mol Cell Biol **27**(16): 5630-5638.

Jurica, M. S. and M. J. Moore (2003). "Pre-mRNA splicing: awash in a sea of proteins." Mol Cell **12**(1): 5-14.

Kadlec, J., D. Guilligay, et al. (2006). "Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1." RNA **12**(10): 1817-1824.

Kadlec, J., E. Izaurralde, et al. (2004). "The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3." Nat Struct Mol Biol **11**(4): 330-337.

Karam, R., J. Wengrod, et al. (2013). "Regulation of nonsense-mediated mRNA decay: Implications for physiology and disease." Biochim Biophys Acta **1829**(6-7): 624-633.

Kawashima, T., M. Pellegrini, et al. (2009). "Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations." RNA **15**(12): 2236-2247.

Kebaara, B., T. Nazarenus, et al. (2003). "The Upf-dependent decay of wild-type PPR1 mRNA depends on its 5'-UTR and first 92 ORF nucleotides." Nucleic Acids Res **31**(12): 3157-3165.

Kebaara, B. W. and A. L. Atkin (2009). "Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae." Nucleic Acids Res **37**(9): 2771-2778.

Keeling, K. M., J. Lanier, et al. (2004). "Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in S. cerevisiae." RNA **10**(4): 691-703.

Kent, O. A., D. B. Ritchie, et al. (2005). "Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway." Mol Cell Biol **25**(1): 233-240.

Kent, W. J. (2002). "BLAT--the BLAST-like alignment tool." Genome Res **12**(4): 656-664.

Kervestin, S. and A. Jacobson (2012). "NMD: a multifaceted response to premature translational termination." Nat Rev Mol Cell Biol **13**(11): 700-712.

Kervestin, S., C. Li, et al. (2012). "Testing the faux-UTR model for NMD: analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p." Biochimie **94**(7): 1560-1571.

Kuperwasser, N., S. Brogna, et al. (2004). "Nonsense-mediated decay does not occur within the yeast nucleus." RNA **10**(12): 1907-1915.

Kuroha, K., T. Tatematsu, et al. (2009). "Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome." EMBO Rep **10**(11): 1265-1271.

Kurosaki, T. and L. E. Maquat (2013). "Rules that govern UPF1 binding to mRNA 3' UTRs." Proc Natl Acad Sci U S A **110**(9): 3357-3362.

Lareau, L. F., M. Inada, et al. (2007). "Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements." Nature **446**(7138): 926-929.

Lee, B. S. and M. R. Culbertson (1995). "Identification of an additional gene required for eukaryotic nonsense mRNA turnover." Proc Natl Acad Sci U S A **92**(22): 10354-10358.

Leeds, P., S. W. Peltz, et al. (1991). "The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor." Genes Dev **5**(12A): 2303-2314.

Lelivelt, M. J. and M. R. Culbertson (1999). "Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome." Mol Cell Biol **19**(10): 6710-6719.

Li, B., C. R. Nierras, et al. (1999). "Transcriptional elements involved in the repression of ribosomal protein synthesis." Mol Cell Biol **19**(8): 5393-5404.

Li, Z., A. G. Paulovich, et al. (1995). "Feedback inhibition of the yeast ribosomal protein gene CRY2 is mediated by the nucleotide sequence and secondary structure of CRY2 pre-mRNA." Mol Cell Biol **15**(11): 6454-6464.

Liao, X. C., J. Tang, et al. (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**(3): 419-428.

Losson, R. and F. Lacroute (1979). "Interference of nonsense mutations with eukaryotic messenger RNA stability." Proc Natl Acad Sci U S A **76**(10): 5134-5137.

Luke, B., C. M. Azzalin, et al. (2007). "Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay." Nucleic Acids Res **35**(22): 7688-7697.

Lund, M. and J. Kjems (2002). "Defining a 5' splice site by functional selection in the presence and absence of U1 snRNA 5' end." RNA **8**(2): 166-179.

Maderazo, A. B., J. P. Belk, et al. (2003). "Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution." Mol Cell Biol **23**(3): 842-851.

Maderazo, A. B., F. He, et al. (2000). "Upf1p control of nonsense mRNA translation is regulated by Nmd2p and Upf3p." Mol Cell Biol **20**(13): 4591-4603.

Maquat, L. E. and G. Serin (2001). "Nonsense-mediated mRNA decay: insights into mechanism from the cellular abundance of human Upf1, Upf2, Upf3, and Upf3X proteins." Cold Spring Harb Symp Quant Biol **66**: 313-320.

Maquat, L. E., W. Y. Tarn, et al. (2010). "The pioneer round of translation: features and functions." Cell **142**(3): 368-374.

Marshall, A. N., M. C. Montealegre, et al. (2013). "Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes." PLoS Genet **9**(3): e1003376.

Messenguy, F., F. Vierendeels, et al. (2002). "Role of RNA surveillance proteins Upf1/CpaR, Upf2 and Upf3 in the translational regulation of yeast CPA1 gene." Curr Genet **41**(4): 224-231.

Meyer, M., M. Plass, et al. (2011). "Deciphering 3'ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing." Mol Cell **43**(6): 1033-1039.

Min, E. E., B. Roy, et al. (2013). "Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit." RNA.

Mishra, S. K., T. Ammon, et al. (2011). "Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing." Nature **474**(7350): 173-178.

Mitchell, P., E. Petfalski, et al. (1997). "The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases." Cell **91**(4): 457-466.

Mitchell, P. and D. Tollervey (2003). "An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation." Mol Cell **11**(5): 1405-1413.

Mitrovich, Q. M. and P. Anderson (2000). "Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans." Genes Dev **14**(17): 2173-2184.

Muhlrad, D. and R. Parker (1994). "Premature translational termination triggers mRNA decapping." Nature **370**(6490): 578-581.

Muhlrad, D. and R. Parker (1999). "Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance." RNA **5**(10): 1299-1307.

Natsoulis, G., C. Dollard, et al. (1991). "The products of the SPT10 and SPT21 genes of Saccharomyces cerevisiae increase the amplitude of transcriptional regulation at a large number of unlinked loci." New Biol **3**(12): 1249-1259.

Nazarenus, T., R. Cedarberg, et al. (2005). "Upf1p, a highly conserved protein required for nonsense-mediated mRNA decay, interacts with the nuclear pore proteins Nup100p and Nup116p." Gene **345**(2): 199-212.

Neubauer, G., A. Gottschalk, 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**(2): 385-390.

Ni, J. Z., L. Grate, et al. (2007). "Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay." Genes Dev **21**(6): 708-718.

Nilsen, T. W. (2005). "Spliceosome assembly in yeast: one ChIP at a time?" Nat Struct Mol Biol **12**(7): 571-573.

Ohnishi, T., A. Yamashita, et al. (2003). "Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7." Mol Cell **12**(5): 1187-1200.

Padgett, R. A., P. J. Grabowski, et al. (1986). "Splicing of messenger RNA precursors." Annual review of biochemistry **55**: 1119-1150.

Page, M. F., B. Carr, et al. (1999). "SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast." Mol Cell Biol **19**(9): 5943-5951.

Pal, M., Y. Ishigaki, et al. (2001). "Evidence that phosphorylation of human Upfl protein varies with intracellular location and is mediated by a wortmannin-sensitive and rapamycin-sensitive PI 3-kinase-related kinase signaling pathway." RNA **7**(1): 5-15.

Peltz, S. W., A. H. Brown, et al. (1993). "mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor." Genes Dev **7**(9): 1737-1754.

Peltz, S. W., J. L. Donahue, et al. (1992). "A mutation in the tRNA nucleotidyltransferase gene promotes stabilization of mRNAs in Saccharomyces cerevisiae." Mol Cell Biol **12**(12): 5778-5784.

Plant, E. P., P. Wang, et al. (2004). "A programmed -1 ribosomal frameshift signal can function as a cis-acting mRNA destabilizing element." Nucleic Acids Res **32**(2): 784-790.

Pleiss, J. A., G. B. Whitworth, et al. (2007). "Rapid, transcript-specific changes in splicing in response to environmental stress." Mol Cell **27**(6): 928-937.

Puig, O., A. Gottschalk, et al. (1999). "Interaction of the U1 snRNP with nonconserved intronic sequences affects 5' splice site selection." Genes Dev **13**(5): 569-580.

Rodriguez-Navarro, S., J. C. Igual, et al. (2002). "SRC1: an intron-containing yeast gene involved in sister chromatid segregation." Yeast **19**(1): 43-54.

Ruiz-Echevarria, M. J., C. I. Gonzalez, et al. (1998). "Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA." The EMBO journal **17**(2): 575-589.

Ruiz-Echevarria, M. J. and S. W. Peltz (2000). "The RNA binding protein Pub1 modulates the stability of transcripts containing upstream open reading frames." Cell **101**(7): 741-751.

Ruiz-Echevarria, M. J., J. M. Yasenchak, et al. (1998). "The upf3 protein is a component of the surveillance complex that monitors both translation and mRNA turnover and affects viral propagation." Proc Natl Acad Sci U S A **95**(15): 8721-8726.

Saha, D., S. Banerjee, et al. (2012). "Context dependent splicing functions of Bud31/Ycr063w define its role in budding and cell cycle progression." Biochem Biophys Res Commun **424**(3): 579-585.

Sayani, S. and G. F. Chanfreau (2012). "Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in Saccharomyces cerevisiae." RNA **18**(8): 1563-1572.

Sayani, S., M. Janis, et al. (2008). "Widespread impact of nonsense-mediated mRNA decay on the yeast intronome." Mol Cell **31**(3): 360-370.

Schlanger, G. and S. M. Friedman (1973). "Ambiguity in a polypeptide-synthesizing extract from Saccharomyces cerevisiae." J Bacteriol **115**(1): 129-138.

Sharp, P. A. (1994). "Split genes and RNA splicing." Cell **77**(6): 805-815.

Sharp, P. A. (2005). "The discovery of split genes and RNA splicing." Trends in biochemical sciences **30**(6): 279-281.

Sheth, U. and R. Parker (2006). "Targeting of aberrant mRNAs to cytoplasmic processing bodies." Cell **125**(6): 1095-1109.

Shirley, R. L., A. S. Ford, et al. (2002). "Nuclear import of Upf3p is mediated by importin-alpha/-beta and export to the cytoplasm is required for a functional nonsense-mediated mRNA decay pathway in yeast." Genetics **161**(4): 1465-1482.

Shirley, R. L., M. J. Lelivelt, et al. (1998). "A factor required for nonsense-mediated mRNA decay in yeast is exported from the nucleus to the cytoplasm by a nuclear export signal sequence." J Cell Sci **111 (Pt 21)**: 3129-3143.

Shirley, R. L., M. R. Richards, et al. (2004). "Using the cre-lox recombination system to assess functional impairment caused by amino acid substitutions in yeast proteins." Biol Proced Online **6**: 209-219.

Strawn, L. A., T. Shen, et al. (2001). "The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex." J Biol Chem **276**(9): 6445-6452.

Suntharalingam, M. and S. R. Wente (2003). "Peering through the pore: nuclear pore complex structure, assembly, and function." Dev Cell **4**(6): 775-789.

Swisher, K. D. and R. Parker (2011). "Interactions between Upf1 and the decapping factors Edc3 and Pat1 in Saccharomyces cerevisiae." PLoS One **6**(10): e26547.

Szer, W. and S. Ochoa (1964). "COMPLEXING ABILITY AND CODING PROPERTIES OF SYNTHETIC POLYNUCLEOTIDES." J Mol Biol **8**: 823-834.

Takahashi, S., Y. Araki, et al. (2008). "Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast." RNA **14**(9): 1950-1958.

Takahashi, S., Y. Araki, et al. (2003). "Interaction between Ski7p and Upf1p is required for nonsense-mediated 3'-to-5' mRNA decay in yeast." The EMBO journal **22**(15): 3951-3959.

Uemura, H. and Y. Jigami (1995). "Mutations in GCR1, a transcriptional activator of Saccharomyces cerevisiae glycolytic genes, function as suppressors of gcr2 mutations." Genetics **139**(2): 511-521.

Umen, J. G. and C. Guthrie (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**(6): 584-597.

Unterholzner, L. and E. Izaurralde (2004). "SMG7 acts as a molecular link between mRNA surveillance and mRNA decay." Mol Cell **16**(4): 587-596.

Vijayraghavan, U., M. Company, et al. (1989). "Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae." Genes Dev **3**(8): 1206-1216.

Vilela, C., B. Linz, et al. (1998). "The yeast transcription factor genes YAP1 and YAP2 are subject to differential control at the levels of both translation and mRNA stability." Nucleic Acids Res **26**(5): 1150-1159.

Vilela, C., C. V. Ramirez, et al. (1999). "Post-termination ribosome interactions with the 5'UTR modulate yeast mRNA stability." The EMBO journal **18**(11): 3139-3152.

Villa, T. and C. Guthrie (2005). "The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing." Genes Dev **19**(16): 1894-1904.

Vogel, J. L., D. A. Parsell, et al. (1995). "Heat-shock proteins Hsp104 and Hsp70 reactivate mRNA splicing after heat inactivation." Curr Biol **5**(3): 306-317.

von der Haar, T. (2008). "A quantitative estimation of the global translational activity in logarithmically growing yeast cells." BMC Syst Biol **2**: 87.

Wahl, M. C., C. L. Will, et al. (2009). "The spliceosome: design principles of a dynamic RNP machine." Cell **136**(4): 701-718.

Wang, W., I. J. Cajigas, et al. (2006). "Role for Upf2p phosphorylation in Saccharomyces cerevisiae nonsense-mediated mRNA decay." Mol Cell Biol **26**(9): 3390-3400.

Wang, W., K. Czaplinski, et al. (2001). "The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts." The EMBO journal **20**(4): 880-890.

Welch, E. M. and A. Jacobson (1999). "An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA." The EMBO journal **18**(21): 6134-6145.

Weng, Y., K. Czaplinski, et al. (1996). "Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein." Mol Cell Biol **16**(10): 5477-5490.

Weng, Y., K. Czaplinski, et al. (1996). "Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover." Mol Cell Biol **16**(10): 5491-5506.

Wente, S. R. and G. Blobel (1993). "A temperature-sensitive NUP116 null mutant forms a nuclear envelope seal over the yeast nuclear pore complex thereby blocking nucleocytoplasmic traffic." J Cell Biol **123**(2): 275-284.

Werner, M., A. Feller, et al. (1987). "The leader peptide of yeast gene CPA1 is essential for the translational repression of its expression." Cell **49**(6): 805-813.

Wilkinson, M. F. (2003). "The cycle of nonsense." Mol Cell **12**(5): 1059-1061.

Yamashita, A., N. Izumi, et al. (2009). "SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay." Genes Dev **23**(9): 1091-1105.

Yamashita, A., T. Ohnishi, et al. (2001). "Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay." Genes Dev **15**(17): 2215-2228.

Yan, B. C., B. A. Westfall, et al. (2001). "Ynl038wp (Gpi15p) is the Saccharomyces cerevisiae homologue of human Pig-Hp and participates in the first step in glycosylphosphatidylinositol assembly." Yeast **18**(15): 1383-1389.

Yost, H. J. and S. Lindquist (1991). "Heat shock proteins affect RNA processing during the heat shock response of Saccharomyces cerevisiae." Mol Cell Biol **11**(2): 1062-1068.

Yun, D. F., T. M. Laz, et al. (1996). "mRNA sequences influencing translation and the selection of AUG initiator codons in the yeast Saccharomyces cerevisiae." Mol Microbiol **19**(6): 1225-1239.

Zhang, S., E. M. Welch, et al. (1997). "Polysome-associated mRNAs are substrates for the nonsense-mediated mRNA decay pathway in Saccharomyces cerevisiae." RNA **3**(3): 234-244.

Zhang, X. and B. Schwer (1997). "Functional and physical interaction between the yeast splicing factors Slu7 and Prp18." Nucleic Acids Res **25**(11): 2146-2152.

Zheng, W., J. S. Finkel, et al. (2008). "Nonsense-mediated decay of ash1 nonsense transcripts in Saccharomyces cerevisiae." Genetics **180**(3): 1391-1405.

Zuk, D. and A. Jacobson (1998). "A single amino acid substitution in yeast eIF-5A results in mRNA stabilization." The EMBO journal **17**(10): 2914-2925.

CHAPTER 2

Nonsense-mediated mRNA Decay Mutes the

Splicing Defects of Splicing Component Mutations

Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations

TADASHI KAWASHIMA,[1] MATTEO PELLEGRINI,[2] and GUILLAUME F. CHANFREAU[1]

[1]Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California at Los Angeles,
Los Angeles, California 90095-1569, USA
[2]Department of Molecular, Cellular and Developmental Biology, University of California at Los Angeles, Los Angeles,
California 90095-1606, USA

ABSTRACT

The role of many splicing factors in pre-mRNA splicing and the involvement of these factors in the processing of specific transcripts have often been defined through the analysis of loss-of-function mutants in vivo. Here we show that inactivating the nonsense-mediated mRNA decay (NMD) results in an enhancement of splicing phenotypes associated with several *S. cerevisiae* splicing factor mutations. Tiling microarrays showed that inactivation of the NMD factor Upf1p in the *prp17Δ* and *prp18Δ* mutant strains results in a larger spectrum of splicing defects than what is observed in the single mutants, including new transcripts previously shown unaffected by Prp17p or Prp18p inactivation. Inactivation of Upf1p in the second step/recycling factor *prp22-1* mutant and in the *nam8Δ* and *mud1Δ* U1 snRNP component mutants also increase unspliced precursor accumulation of several specific transcripts. In addition, deletion of *UPF1* partially suppresses the growth defects associated with the *prp17Δ* or *prp22-1* mutations, demonstrating a positive genetic interaction between NMD and splicing factor mutants. These results show that RNA surveillance by NMD can mask some of the effects of splicing factor mutations, and that the roles of splicing factors cannot be fully understood in vivo unless RNA degradation systems that degrade unspliced precursors are also inactivated.

Keywords: NMD; splicing; yeast; RNA degradation; quality control

INTRODUCTION

Nonsense-mediated mRNA decay (NMD) is an RNA degradation mechanism that degrades pre-mRNA containing premature translation termination codons (PTCs) (Behm-Ansmant et al. 2007; Chang et al. 2007; Isken and Maquat 2008). Due to the high probability of encountering a PTC, unspliced pre-mRNAs or some alternatively spliced RNAs are potential targets for NMD. Indeed, NMD is involved in regulating highly conserved alternatively spliced forms containing PTCs (Lareau et al. 2007; Ni et al. 2007) and degrades unspliced precursors of transcripts spliced with suboptimal efficiency in *C. elegans* (Mitrovich and Anderson 2000), *S. cerevisiae* (He et al. 1993; Sayani et al. 2008), and *Paramecium* (Jaillon et al. 2008). These studies led to the idea that NMD might act as a general quality control

mechanism for defective or suboptimal splicing. Based on this, one might expect NMD to be also involved in degrading unspliced precursors generated by mutations of splicing signals or by inactivation of splicing factors. Surprisingly, combining a thermosensitive mutation of the Prp2p splicing factor with NMD inactivation did not result in the stabilization of several unspliced precursors (Bousquet-Antonelli et al. 2000). NMD integrity also does not affect the steady-state levels or the rate of decay of unspliced precursors of splicing substrates containing a mutated actin intron (Hilleren and Parker 2003; Sayani et al. 2008). In addition, several other unspliced precursors resulting from splicing signal mutations accumulate to high levels in the presence of active NMD (Vijayraghavan et al. 1986; Chanfreau et al. 1994; Sayani et al. 2008). These observations led to the conclusion that NMD does not contribute to the quality control of splicing in splicing mutants. However, the recent observation that NMD degrades most unspliced precursors resulting from a 5′-splice site mutation in the *RPS10B* gene (Sayani et al. 2008) led us to investigate whether NMD can mask or reduce the effects of *trans*-acting splicing factor mutations. In this study, we

show that many transcripts do not exhibit any detectable splicing defect in several spliceosome component mutants unless NMD is also disrupted. In addition we also show that the quantitative effects of these splicing mutants are generally enhanced in the absence of functional NMD. These results reveal that NMD can mask the effects of splicing factor inactivation by degrading unspliced RNAs that accumulate due to defective splicing. Our observations show that the full extent of the role of splicing factors cannot be fully understood in vivo unless RNA degradation systems that eliminate these unspliced RNAs are inactivated as well.

RESULTS

NMD quantitatively reduces the splicing defects associated with inactivation of the Prp17p and Prp18p splicing factors

To investigate whether NMD can mask the effects of splicing factor mutations in vivo, we first inactivated the NMD factor Upf1p in yeast strains carrying deletions of the nonessential splicing factors Prp17p and Prp18p. The choice of these nonessential splicing factors was based on the fact that unspliced RNA levels could be analyzed at steady-state and normal growth temperatures without having to shift the cells to non-permissive temperatures, alleviating the issues of viability and mild heat-shock conditions. We constructed the double knockout strains *prp17Δupf1Δ* and *prp18Δupf1Δ* by direct knockout of *UPF1* in the *prp17Δ* or *prp18Δ* backgrounds. We then performed tiling array analysis of RNAs and analyzed the intronic signals in all single and double mutants strains. To investigate the extent to which NMD can mask splicing defects in the *prp17Δ* and *prp18Δ* strains, we generated Z-scores based on measuring the ratio of intronic signals between each strain (provided in Supplemental Table 1). Our previous study had shown that analysis of the variations in intronic signals is a valid first approximation of the amount of unspliced precursor that accumulates in mutant strains compared with the wild type (Sayani et al. 2008). Intronic Z scores were based on the average of the log2 of the ratio of the signals obtained from

probes located in introns or spanning the exon–intron junctions between each of the strains analyzed. Because the number of probes varied depending on introns, the use of the Z-scores allowed us to take into account the number of probes used to measure the signal from each intron. Comparison of the Z-scores between the *prp17Δ* mutant and the wild type (Fig. 1A, x-axis), and between the *prp17Δupf1Δ* and the *upf1Δ* mutant (Fig. 1A, y-axis) showed that the quantitative effect of the *prp17Δ* deletion on intronic signal is much more pronounced when NMD is inactive (Fig. 1A, left panel). This is illustrated by the large number of introns that fall above the diagonal line, or for which the Z-score is close to or lower than zero for the *prp17Δ* mutant to the wild-type comparison (Fig. 1A, x-axis), but which exhibit an increase of Z-score above zero in the double mutant (Fig. 1A, y-axis). The same effect was observed when comparing the effect of NMD inactivation (Fig. 1A, right panel, *upf1Δ* versus WT) to the effects of NMD inactivation in the context of the *prp17Δ* deletion (Fig. 1A, right panel,
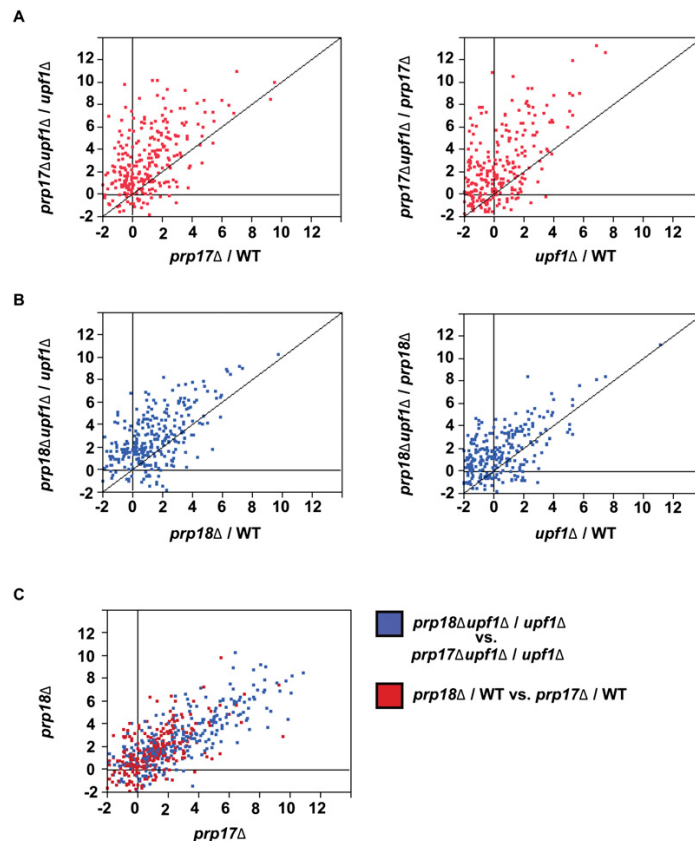

FIGURE 1. Quantitative analysis of intronic signals in the *prp17Δ*, *prp18Δ*, and *upf1Δ* mutants and in double mutants. Plotted are the Z-scores for the average of the log2 of the ratio of intronic signals for all introns of the *S. cerevisiae* genome in the indicated strain comparisons. Lines indicate the diagonal on which equal Z-score for the strain comparisons would be located.

prp17Δupf1Δ versus *prp17Δ*). Overall intronic signals were generally higher in the *prp17Δupf1Δ* double deletion strain than in either single mutant, which shows that NMD quantitatively reduces intronic signals in the *prp17Δ* deletion mutant. The same effect was observed for the *prp18Δ upf1Δ* double mutant when compared with either *prp18Δ* or *upf1Δ* single mutants (Fig. 1B).

To provide an independent way to visualize the effect of the *prp17Δ* and *prp18Δ* mutations in the context of active or inactive NMD, we plotted the intronic scores of the comparison sets of these mutants against each other in the context of active NMD (Fig. 1C, red dots, each single mutant is compared with wild type) or inactive NMD (blue dots; each double mutant is compared with the *upf1Δ* single mutant). When plotted on the same graph, the data points observed in the context of inactive NMD (Fig. 1C, blue dots) are shifted to higher values than those observed when the effects of *prp17Δ* and *prp18Δ* mutants are compared when NMD is active (Fig. 1C, red dots). This observation shows that the deletion of *PRP17* or *PRP18* results in higher intronic signals when these mutations are combined with NMD inactivation. We note that the effects of Prp17p and Prp18p inactivation are generally well correlated, whether NMD is active or not (Fig. 1C), with only a small minority of introns showing differential effects upon inactivation of Prp17p or Prp18p. Overall, these data show that NMD reduces intronic signal accumulation in splicing factor mutants when analyzing the entire intronic population of *S. cerevisiae*. Even if a fraction of the intronic signal comes from lariat intermediates generated by inactivation of these second step splicing factors, these results suggest that a significant fraction of the unspliced precursors generated by spliceosome component mutations are degraded by NMD, which limits or prevents their accumulation in vivo.

NMD mutes transcript-specific splicing defects associated with inactivation of Prp17p and Prp18p

The previous analysis revealed that inactivation of Upf1p in the *prp17Δ* or *prp18Δ* mutants can enhance the intronic signal accumulation phenotype associated with these splicing mutants, suggesting that the splicing defects of these mutants can be more readily detected or is exacerbated when NMD is inactivated. We next sought to identify transcripts for which NMD would completely mask the effects of the Prp17p or Prp18p depletion. Such transcripts would exhibit no increase of intronic signal in the *prp17Δ* and *prp18Δ* mutants compared with wild type, but a large signal increase in the double mutants when NMD is also inactivated (Fig. 1A,B, upper left quadrants). To facilitate visualization of these transcripts, we performed hierarchical clustering analysis based on the intronic Z-scores and searched for transcripts with no intronic signal accumulation in the *prp17Δ* or *prp18Δ* versus wild-type comparison,

but for which an increase of intronic signal was detected in the double-mutant strains when compared with each single mutant (Figs. 2, 3). This analysis revealed several transcripts for which intronic signal increase is not detectable in either of the single mutants, but for which a strong increase in intronic signal can be detected when comparing *prp17Δupf1Δ* to *upf1Δ* or *prp17Δupf1Δ* to *prp17Δ* (Fig. 2). Similar observations were made for the *prp18Δ*-derived mutants, although the number of introns affected was smaller than for *prp17Δ* (Fig. 3). The intronic signal increase observed for many introns does not result from a general splicing defect in the double mutants, since several precursors did not show an increase of intronic signal in either the single or double mutants (Figs. 2, 3, Clusters A). The lack of effect observed for these transcripts was not due to the absence of premature termination codons (PTCs), as the large majority of yeast unspliced pre-mRNAs contain PTCs, as shown previously (Sayani et al. 2008). Interestingly, this analysis showed that the absence of Prp17p can result in intronic signal increase for short introns, in contrast to what previous analyses had shown (Clark et al. 2002; Sapra et al. 2004). For example, the *BOS1, CIN2,* and *MTR2* genes showed an increase of intronic signal only in the context of the *prp17Δupf1Δ* double mutant, but not in the *prp17Δ* or *upf1Δ* single mutants (Fig. 2, Cluster B). The small sizes of the *BOS1, CIN2,* and *MTR2* introns (73, 80, and 99 nucleotides [nt], respectively) show that Prp17p can affect the splicing of introns shorter than 100 nt, but that this effect is masked by NMD. We also observed a strong increase of intronic signal in double mutants for a large number of ribosomal protein genes (Figs. 2, 3, Clusters B,C,E,F). These genes were not found to accumulate unspliced precursors in NMD mutants in a normal splicing background (Sayani et al. 2008). These results show that NMD can affect a larger number of transcripts than previously estimated. However, in the absence of splicing mutations, the splicing of these transcripts is robust enough to prevent escape of unspliced precursors from the spliceosome.

The previous analyses measured intronic signal accumulation; therefore, it was possible that the increase of signal observed in the double mutants may be indicative of accumulation of RNA species other than the unspliced precursors, such as lariat intermediates that may result from the inactivation of the Prp17p and Prp18p second step splicing factors. To investigate whether or not the increase of signal observed was due to unspliced precursor stabilization, we first performed Northern blot analysis on the *RPL16A, RPS13,* and *RPS11B* transcripts (Fig. 4A). Based on the tiling array measurements, *RPL16A* and *RPS11B* were predicted to have stronger precursor accumulation in the double mutants than each of the corresponding single mutants, while unspliced precursor accumulation for *RPS13* was predicted to be found in the *upf1Δ* strain and not exacerbated by deletion of Prp17p or

FIGURE 2. Clustering analysis of intronic Z-scores for the *prp17Δ*, *upf1Δ*, and *prp17Δupf1Δ* mutants when compared with wild-type or single mutants.

Prp18p. Northern analysis validated the prediction made by the tiling array analysis, with a dramatic accumulation of slower migrating species in the *prp17Δupf1Δ* or *prp18Δupf1Δ* mutants compared with the single mutants for *RPL16A* and *RPS11B*. This was not observed for *RPS13*, for which the accumulation of the slower migrating species was not increased by disruption of Prp17p or Prp18p. We also observed slightly faster migrating species that are detectable in the *prp17Δ* or *prp18Δ* single mutants that might correspond to intron-exon2 lariat intermediates
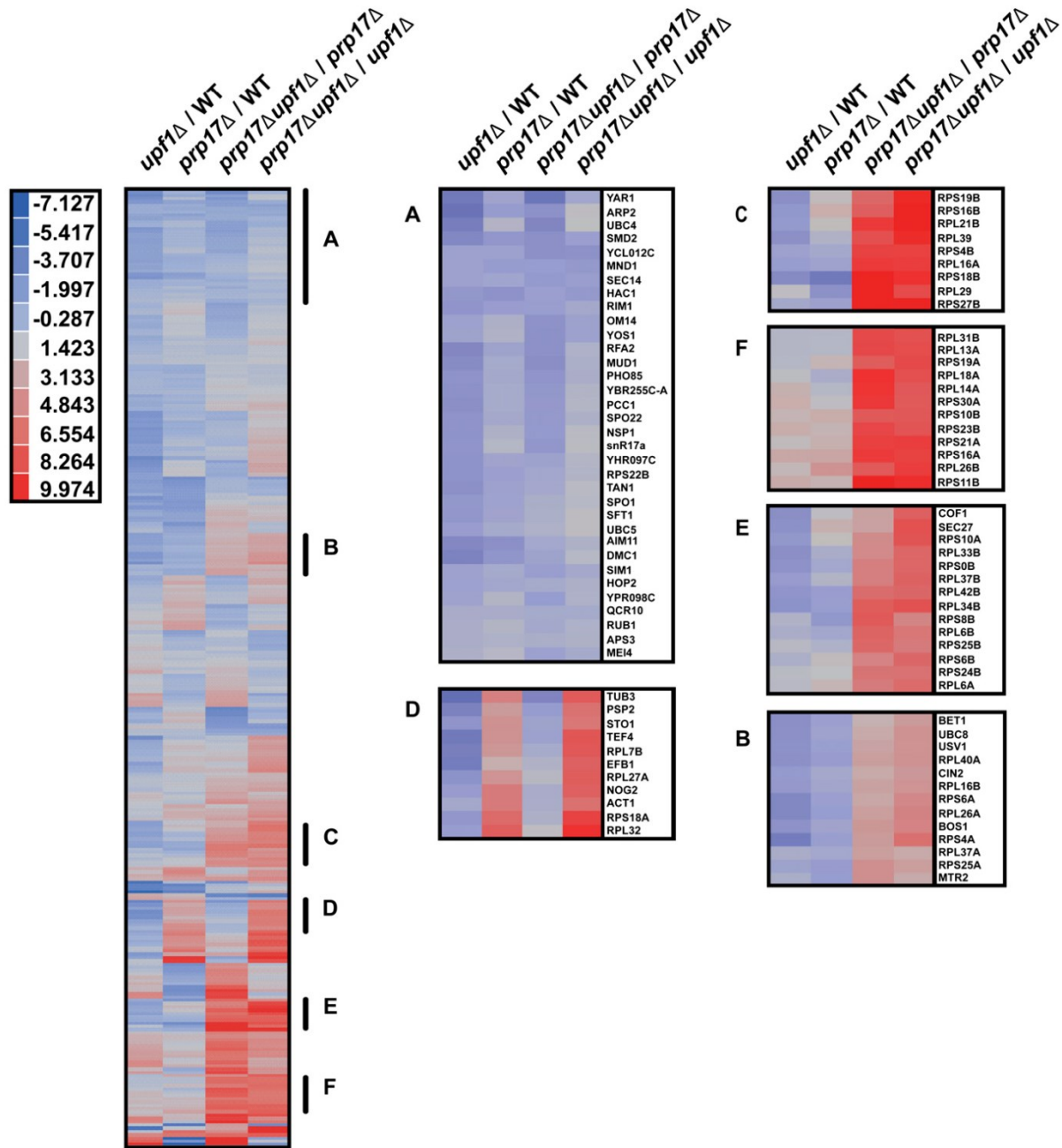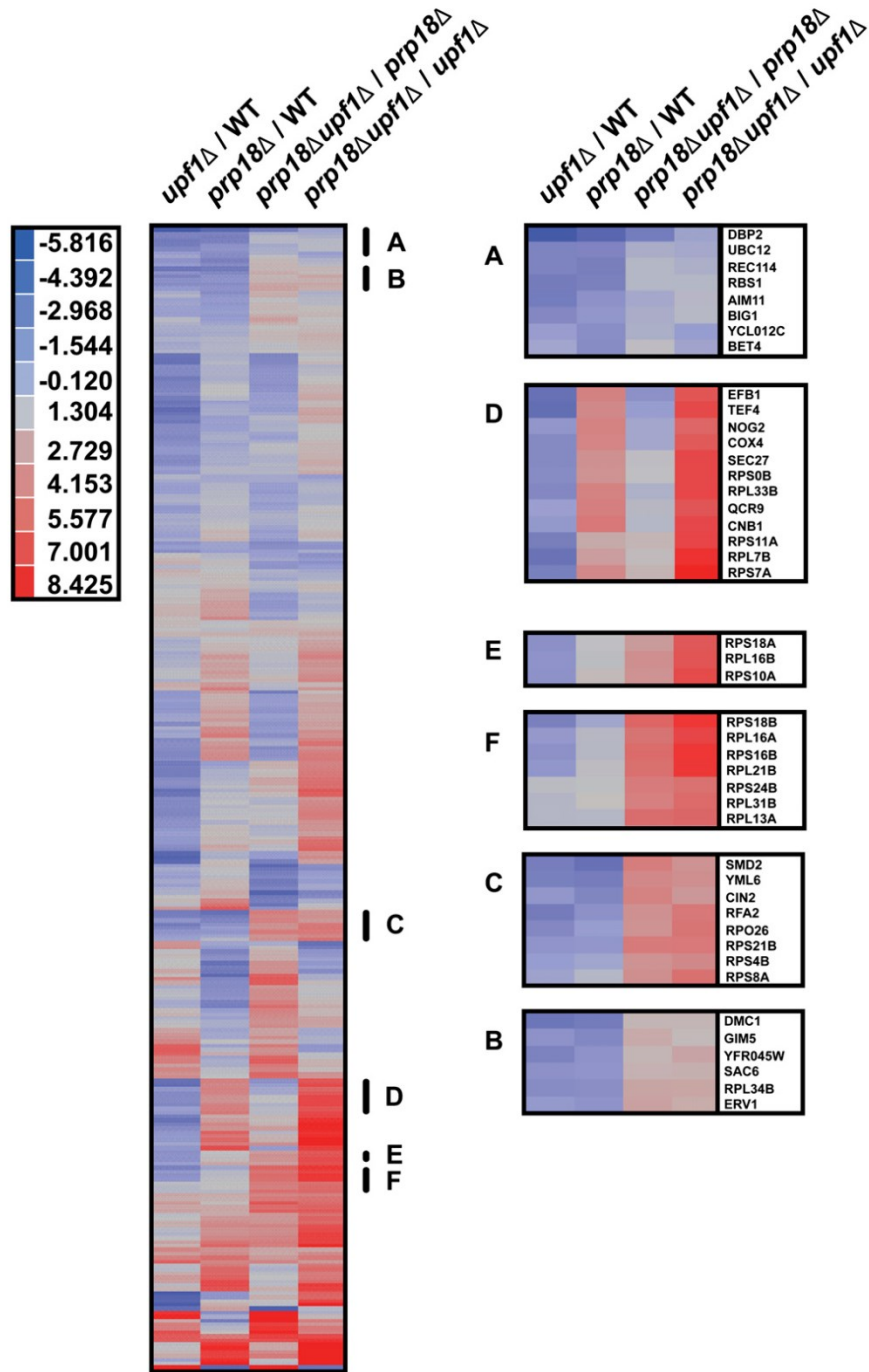
FIGURE 3. Clustering analysis of intronic Z-scores for the *prp18Δ*, *upf1Δ*, and *prp18Δupf1Δ* mutants when compared with wild-type or single mutants.
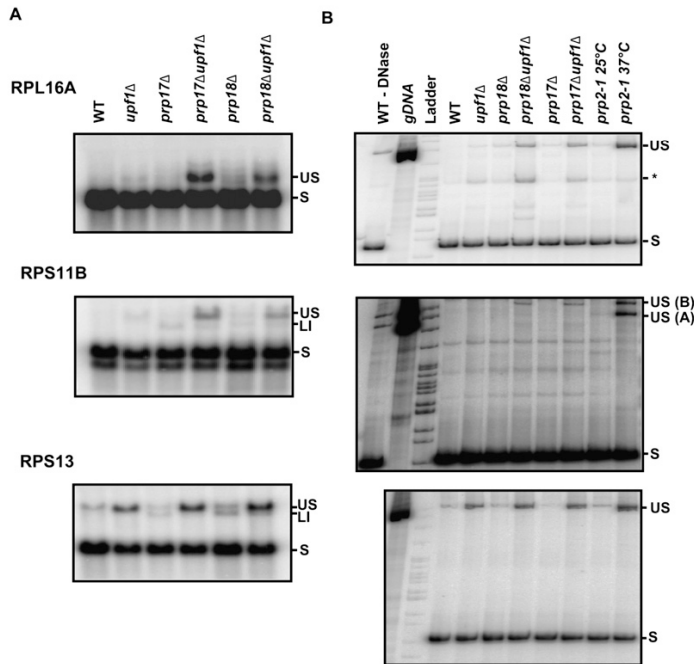
FIGURE 4. Northern blot and RT-PCR analysis of splicing in wild-type, *upf1Δ*, *prp17Δ*, *prp18Δ*, *prp17Δupf1Δ*, and *prp18Δupf1Δ* mutants. Shown are Northern blots (*A*) or RT-PCR (*B*) analysis of the indicated mRNAs (*RPL16A, RPS11B, RPS13*) in the corresponding strains. For the RT-PCR analysis, a genomic DNA sample and RNAs extracted from a *prp2-1* strain grown at 25°C or shifted for 30 min to 37°C were included as positive controls for the detection of unspliced species. The migration of the unspliced pre-mRNAs and spliced mRNAs are indicated by US and S, respectively. LI indicates the lariat intron-exon2 intermediates detected by Northern blot in the *prp17Δ, prp18Δ* mutants. The species labeled * for *RPL16A* were not mapped and might correspond to a cryptic splicing event. For the *RPS11* RT-PCR panel, the oligonucleotides used did not discriminate between the *RPS11A* and *RPS11B* copies; therefore, both unspliced species are amplified by PCR.

positive controls for the accumulation of unspliced precursors, as well as a PCR reaction performed on genomic DNA (gDNA) or on wild-type RNAs that were not treated with DNase (Fig. 4B, first two lanes for *RPL16A* and *RPS11B*). This analysis showed that unspliced (US) precursors of *RPL16A* and *RPS11B* accumulate specifically in the *prp17Δupf1Δ* or *prp18Δupf1Δ* double mutants, in contrast to single mutants in which these species are barely visible or undetectable. Thus, the slower migrating species detected on Northern blots in Figure 4A, those that accumulate specifically in the double mutants, correspond to unspliced precursors. We note that the accumulation of unspliced precursor in the *prp17Δupf1Δ* or *prp18Δupf1Δ* double mutants is specific to *RPS11B*, as unspliced species of *RPS11A* cannot be detected in these strains. This is in contrast with the *prp2-1* mutant, in which both *RPS11A* and *RPS11B* unspliced precursors can be detected (Fig. 4B, middle panel). Finally, analysis of *RPS13* by RT-PCR confirmed the observation made by Northern blot that the accumulation of unspliced precursors of *RPS13* in the *upf1Δ* mutant is not exacerbated by inactivation of Prp17p or Prp18p (Fig. 4B, bottom panel).

(Fig. 4A, LI). However, their accumulation is not exacerbated by Upf1p inactivation in contrast to the presumed unspliced precursors (Fig. 4A). This is in agreement with previous studies showing that lariat intermediates are degraded in the cytoplasm in a Dbr1p-dependent manner (Hilleren and Parker 2003). This Northern analysis suggested that unspliced precursor accumulation of *RPS11B* and *RPL16A* was exacerbated in the *prp17Δupf1Δ* or *prp18ΔupflΔ* double mutants compared with the single mutants. However, due to the small size of exon1, we could not rule out that some of these species postulated to be unspliced precursors may actually correspond to lariat intermediates. To further investigate this possibility we performed RT-PCR on DNase-treated RNA samples from these strains using oligonucleotides hybridizing to exon1 and exon2 sequences. These oligonucleotides specifically amplify mRNAs and unspliced precursors, allowing their detection independently from lariat intermediates (Fig. 4B). We also included in this analysis samples extracted from the *prp2-1* splicing mutant grown at 25°C or shifted for 30 min at 37°C (Fig. 4B, last two lanes) to serve as

RT-PCR analyses show that the effect of NMD in masking the effects of splicing defects is more widespread than revealed by tiling arrays

Even when considering the RT-PCR data described above, some of the intronic signals detected by tiling arrays may result either from unspliced precursors or from lariat species such as intron-exon2 intermediates that can accumulate in the *prp17Δ* or *prp18Δ* second step mutants. For some of the intron-containing genes, we considered the possibility that we did not detect any increase of signal in the *prp17ΔupflΔ* or *prp18ΔupflΔ* double mutants compared with the single splicing mutants, because the signal coming from lariat intermediates that accumulate in the second-step *prp17Δ* or *prp18Δ* mutants exceeds the signal resulting from unspliced precursors in the double mutants. To extend our studies, we analyzed by RT-PCR unspliced precursor accumulation for five additional transcripts, *TUB1, TAF14, LSB3, ACT1,* and *MER2,* which exhibited a wide range of intronic Z-scores as determined by the tiling arrays

(Supplemental Table 1). Based on the Z-scores, the *prp17Δupf1Δ* or *prp18Δupf1Δ* double mutants were not predicted to exhibit higher unspliced precursor accumulation than the single *prp17Δ* or *prp18Δ* mutants for *TUB1*, *LSB3*, and *TAF14*. However RT-PCR analysis showed that unspliced precursor accumulation is exacerbated for these three genes in the *prp18Δupf1Δ* and *prp17Δupf1Δ* double mutants compared with each single mutant (Fig. 5). In contrast, actin (*ACT1*) precursors were undetectable in both the *prp17Δupf1Δ* and *prp18Δupf1Δ* double mutants, while these species could be detected in the *prp2-1* mutant. In addition, a meiotic transcript which naturally accumulates unspliced pre-mRNAs in vegetative conditions such as *MER2* was also unaffected by NMD inactivation, either alone or in combination with splicing mutants (Fig. 5). This result corroborates previous observations that unspliced precursors of meiotic transcripts are not subject to NMD, possibly because of nuclear retention (Scherrer and Spingola 2006). We conclude that the tiling array analysis provided us with a minimal list of genes for which unspliced precursor accumulation is exacerbated by inactivation of NMD in the *prp17Δ* or *prp18Δ* mutants; however, the effects are likely to be more widespread than found using the arrays, because the intronic signal that results from lariat intermediate accumulation in the *prp17Δ* or *prp18Δ* mutants prevents an accurate measure-ment of the accumulation of unspliced precursors for a subset of intron-containing transcripts.

NMD reduces the amount of several unspliced precursors in the *prp22-1*, *nam8Δ*, and *mud1Δ* splicing mutants

The previous analysis showed that inactivation of Upf1p can enhance or reveal unspliced precursor accumulation phenotypes associated with deletion of the spliceosome components Prp17p or Prp18p by repressing the degradation of some of the precursors normally targeted by NMD. To extend these results to other splicing mutants, we combined the *prp22-1* thermosensitive mutant (Vijayraghavan et al. 1989) and the U1 snRNP components *nam8Δ* (Gottschalk et al. 1998) or *mud1Δ* (Liao et al. 1993) deletion mutants with the *upf1Δ* deletion by direct knock-out of *UPF1* in the corresponding single mutants. We grew the *prp22-1 upf1Δ* mutant strain and corresponding single mutants at 25 or 30°C to investigate whether inactivation of Upf1p could reveal or enhance some splicing defects in the *prp22-1* mutant, even at permissive temperatures. This approach allowed us to avoid shifting the cells to non-permissive temperature, because such a shift might induce growth arrest in the *prp22-1* mutant, with possible interference with the NM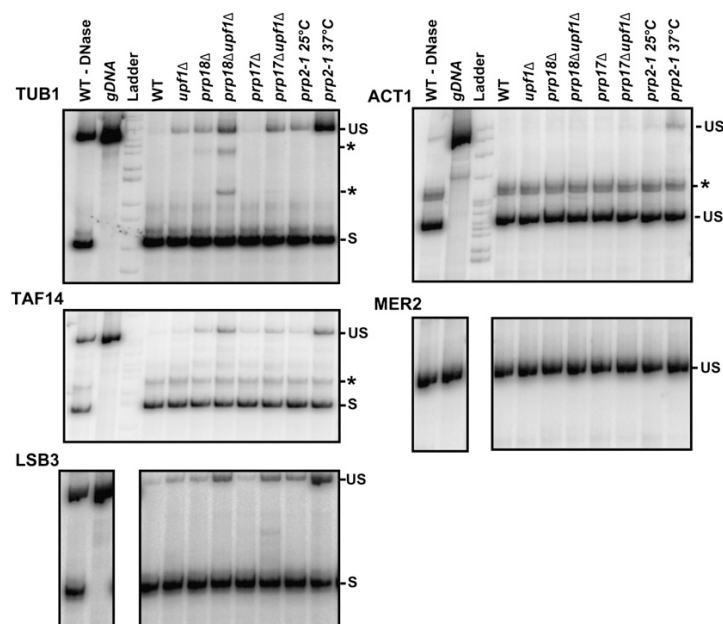D process. We then investigated by RT-PCR the accumulation of unspliced transcripts for the same subset of transcripts previously analyzed for the *prp17Δ* or *prp18Δ* mutants series (Fig. 6). This analysis showed that the level of unspliced *RPL16A*, *RPS11B*, and *TAF14*, precursors detected by RT-PCR was slightly higher in the *prp22-1* mutant grown at 30°C, but was exacerbated when the *upf1Δ* deletion was combined with this mutation. This effect was specific to these transcripts, as it was not observed for the *RPS13*, *LSB3*, *MER2*, *TUB1*, and *ACT1* genes (Fig. 6). In contrast, we did not detect exacerbation of unspliced precursor accumulation when the *prp22-1* and *prp22-1upf1Δ* mutants were grown at 25°C (Fig. 6). It is possible that the NMD pathway is less active at this lower temperature, or that the 30°C temperature used previously exacerbates the degradation of unspliced precursors that are generated in the *prp22-1* mutant.

In the case of the *mud1Δ* and *nam8Δ* mutants for which the corresponding proteins are part of the U1 snRNP and are involved in the first step of splicing (Liao et al. 1993; Gottschalk et al. 1998), we found similarly that inactivation of

FIGURE 5. RT-PCR analysis of splicing of the *TUB1*, *TAF14*, *LSB3*, *ACT1*, and *MER2* genes in wild-type, *upf1Δ*, *prp17Δ*, *prp18Δ*, *prp17Δupf1Δ*, and *prp18Δupf1Δ* mutants. The species labeled * for *TAF14* and *ACT1* were not mapped and might correspond to cross-hybridization of the primers to another transcript, since these are detected regardless of the strains studied. The species labeled * for *TUB1* were not mapped and might correspond to a cryptic splicing event.
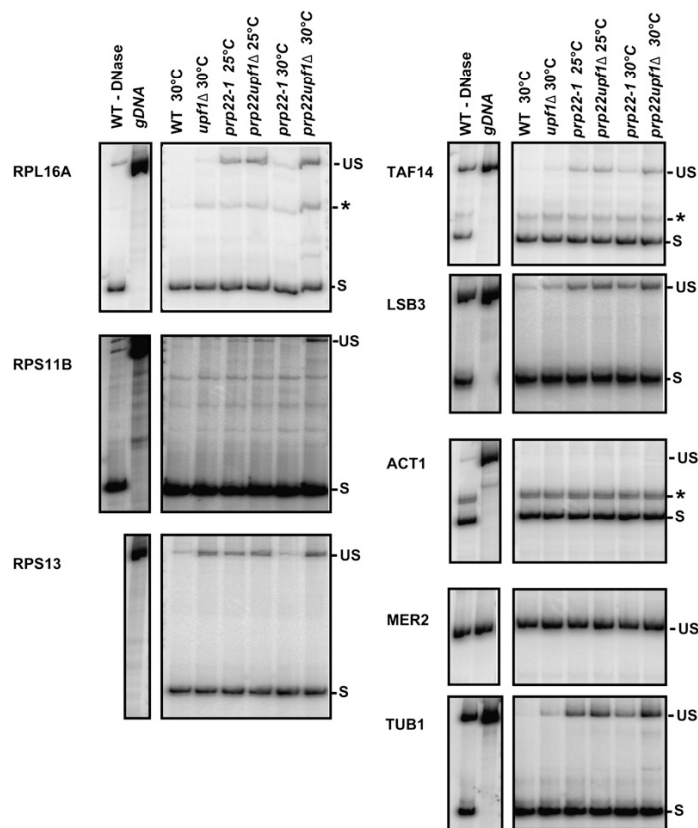
FIGURE 6. RT-PCR analysis of splicing in wild-type, *upf1Δ*, *prp22-1*, and *prp22-1 upf1Δ* mutants; legend as in Figures 4 and 5.

Upf1p in these mutants increased the accumulation of unspliced forms of *RPL16A* and *RPS11A* (Fig. 7). In addition, *TAF14* unspliced pre-mRNAs were observed only in the *mud1Δupf1Δ* double mutant, showing that inactivating the NMD can reveal splicing defects specific to the deletion of *MUD1*, which are not observed when Nam8p is inactivated. In contrast, inactivation of Upf1p did not reveal nor exacerbate the accumulation of unspliced precursors for the *LSB3*, *ACT1*, *MER2*, and *TUB1* transcripts. As observed for the previous mutants, inactivation of Mud1p or of Nam8p had little effect on the accumulation of unspliced *RPS13* transcripts, which are already abundant in the *upf1Δ* mutant. Thus, similarly to what we found for the second step mutants *prp17Δ*, *prp18Δ*, and *prp22-1*, inactivation of Upf1p can reveal some of the splicing defects that are otherwise masked by NMD in first step mutants.

Genetic interactions between *upf1Δ* and splicing mutants

Following the previous observations, we could not exclude the possibility that the increase of unspliced precursors

observed in the double mutants might have been caused by synergistic growth defects. While generating the double mutants combining the *upf1Δ* deletion with the splicing mutants, we found, to the contrary, that the double mutants in which the *prp17Δ* or *prp22-1* mutation are combined with the *upf1Δ* deletion exhibited a better growth rate than those of the single *prp17Δ* or *prp22-1* mutant (Fig. 8). Inactivating Upf1p was, however, not sufficient to rescue the temperature-sensitive phenotype of the *prp22-1* mutant (data not shown). Therefore, inactivation of NMD partially suppresses the growth defects associated with the *prp17Δ* or *prp22*-1 splicing mutants. This is in contrast to the deletion of Prp18p, which shows a synthetic growth defect when combined with deletion of Upf1p (Fig. 8). This result is reminiscent of that observed for the splicing/nuclear retention factor BBP/ScSF1, for which inactivation of NMD results in a synthetic lethal phenotype (Rutz and Seraphin 2000). The reduced growth of the *prp18Δupf1Δ* double mutant compared with the *prp18Δ* mutant might explain why the quantitative effects of NMD inactivation were less pronounced for the *prp18Δ* mutant than for the *prp17Δ* mutant (Figs. 1–3).

We also monitored the growth rates of *nam8Δ*, *mud1Δ*, *upf1Δ*, and corresponding double mutants, but neither of the single or double mutants exhibited a significant growth defect on plates (data not shown). Thus, the positive genetic interactions observed between the *upf1Δ* strain and splicing mutants seem to be specific to *prp22-1* and *prp17Δ*. The growth suppression phenotypes observed between the *prp22-1* and *upf1Δ* deletion and *prp17Δ* and *upf1Δ* are rather surprising, given the fact that previous work had shown that the combination of a mutation of the splicing/ retention factor BBPS/ScSF1 and NMD inactivation results in a synthetic lethal phenotype (Rutz and Seraphin 2000) and given the negative genetic interaction observed between the *upf1Δ* deletion and the *prp18Δ* mutant (Fig. 8). However, this result can be explained by considering the splicing of a specific group of transcripts that might be rate limiting for growth. It is possible that the absence of Prp17p or the partial inactivation of Prp22p at permissive temperature results in reduced production of these rate-limiting spliced mRNAs. Inactivation of NMD in this context might increase the level of the corresponding unspliced RNAs, which might ultimately increase the production of spliced
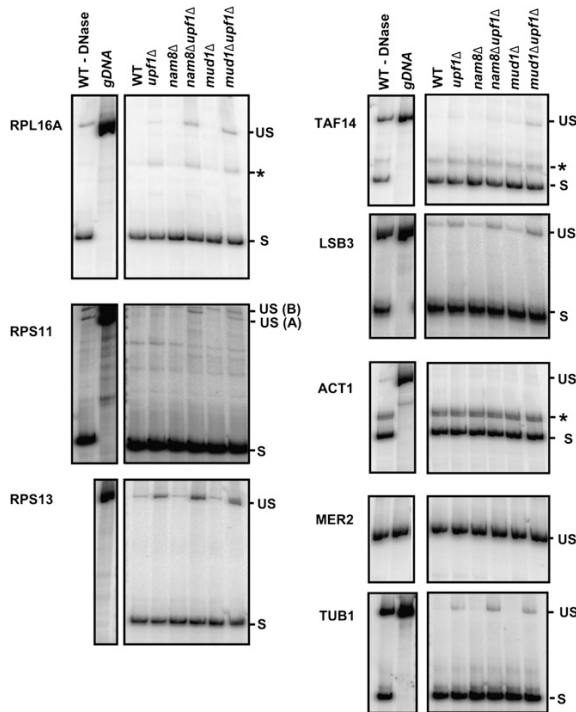
FIGURE 7. RT-PCR analysis of splicing in wild-type, *upf1Δ*, *nam8Δ*, *mud1Δ*, *nam8Δupf1Δ*, and *mud1 upf1Δ* mutants; legend as in Figures 4 and 5.

mRNAs, because more precursors are available for splicing. A similar model has been proposed to explain the suppression of branchpoint mutations by the Prp16p splicing factor mutations (Burgess et al. 1990). However, we were unable to identify transcripts using tiling arrays for which exonic signal is increased specifically in the *prp17Δ upf1Δ* mutant compared with the *prp17Δ* and *upf1Δ* mutants, which would have indicated increased splicing. Thus, it is possible that the suppression observed is due to other effects (see Discussion). Regardless of the interpretation, these results identify a positive genetic interaction between an NMD mutant and several splicing factor mutants.

DISCUSSION

In this study we have shown that NMD can degrade a fraction of unspliced precursors generated by deletion or inactivation of the spliceosome components Prp17p, Prp18p, Prp22p, Mud1p, and Nam8p. Thus, NMD inactivation enhances the quantitative effects of several splicing factor mutations and also reveals transcripts that were thought to be unaffected by these splicing factors. Interestingly, our analysis reveals that the *prp17Δ* and *prp18Δ* mutants exhibit accumulation of unspliced precursors (Figs. 4, 5). The Prp17p and Prp18p spliceosome

components are thought to be involved specifically in the second step of splicing (Jones et al. 1995; Umen and Guthrie 1995), yet they exhibit a first step defect as shown by the accumulation of unspliced precursors in vivo (Figs. 4, 5). This observation could be explained by a hitherto unidentified involvement of these factors in the first step of splicing. Indeed, Prp17p was found to be associated biochemically with pre-first step C-complex in mammalian cells (Bessonov et al. 2008) and in pre-catalytic spliceosomes in yeast (Sapra et al. 2008). This association with pre-first step complexes might provide an explanation for the first-step defect observed for strains lacking Prp17p in vivo. Alternatively, it is possible that disrupting second-step splicing factors may shift the equilibrium of spliceosome conformations toward a first step specific conformation, as suggested by recent genetic analyses (Konarska et al. 2006; Liu et al. 2007). This equilibrium shift might then favor the reversal of the first step in vivo, resulting in the reverse production of unspliced precursors. This hypothesis, although complex, is supported by the observation that both steps of pre-mRNA splicing are reversible, at least in vitro (Tseng and Cheng 2008). Finally, it is possible that the unspliced precursor phenotype observed in these second-step mutants is indirect and due to the sequestration of a limiting first step splicing factor in spliceosomes stalled prior to the second catalytic step in vivo.

Based on the results we obtained with the *prp17Δ*, *prp18Δ*, *nam8Δ*, *mud1Δ*, and *prp22-1* mutants, it is puzzling that a previous analysis of the *prp2-1* splicing mutant did not result in the stabilization of unspliced pre-mRNAs upon inactivation of NMD (Bousquet-Antonelli et al. 2000). It is possible that unspliced precursors in the *prp2-1* mutant are
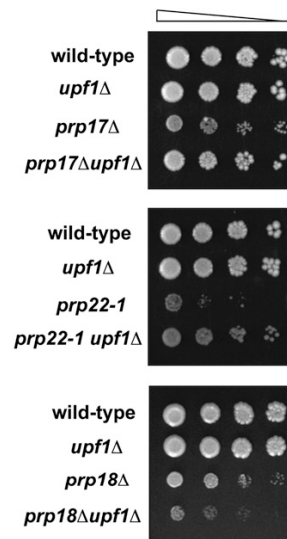

FIGURE 8. Growth of wild-type, *upf1Δ*, splicing mutants and the corresponding double mutants. Plates were incubated for 4 d at 25°C.

trapped in a pre-catalytic conformation of the spliceosome in the nucleus and specifically degraded by the nuclear exosome because of nuclear retention. Alternatively, it is possible that since the analysis of the *prp2-1* splicing mutant was performed on only a few transcripts (Bousquet-Antonelli et al. 2000), some unspliced transcripts generated by the *prp2-1* mutant and degraded by NMD may have escaped identification.

Our results show that NMD targets many more unspliced precursors than we previously estimated (Sayani et al. 2008). In natural conditions, some transcripts are spliced with very high efficiency, and therefore generate very few unspliced precursors. However, in conditions where splicing is suboptimal due to nonessential splicing factors inactivation (*prp17Δ*, *prp18Δ*, *nam8Δ*, and *mud1Δ*), or because of partial inactivation of an essential splicing factor at a permissive temperature (*prp22-1*), unspliced precursors escape the spliceosome and are degraded by NMD. This process might compete with splicing, as shown by the partial rescue of the growth defect of the *prp17Δ* and *prp22-1* mutations by Upf1p inactivation (Fig. 6). We do not yet understand the molecular basis for this suppression and why the positive genetic interactions are specific to *prp17Δ* and *prp22-1*. One interpretation is that inactivation of Upf1p might increase the growth of these splicing mutants by increasing the level of unspliced precursors of transcripts that are rate limiting for growth, thereby increasing the pool of splicing substrates for these transcripts and possibly the amount of mature mRNA produced. Since pre-mRNA splicing is nuclear, and NMD is cytoplasmic— or perinuclear—this would require the shuttling back of unspliced pre-mRNAs to the nucleus. Unfortunately, based on inspection of the tiling arrays, we have not been able to identify transcripts for which spliced transcripts are increased in the *prp17Δupf1Δ* double mutant. Therefore, it is formally possible that these genetic interactions rely on molecular interactions that are unrelated to splicing.

Several studies have investigated the effects of spliceosome mutations in a genome-wide manner to identify transcripts whose splicing requires specific spliceosome components (Clark et al. 2002; Sapra et al. 2004; Pleiss et al. 2007). Conclusions about transcript-specific effects of spliceosome component mutations based on the measurement of precursor accumulation are problematic, since the absence of effect of a mutation on the accumulation of specific unspliced transcripts might have been the result of RNA degradation masking the accumulation of unspliced precursors, as shown in this study. Analysis of the reduction of mature mRNA production is a more informative parameter to measure splicing defects directly in vivo. However, in the case of stable mRNAs, splicing defects do not necessarily result in a reduction of mature mRNAs (Pikielny and Rosbash 1985), making this approach impractical in many cases, especially in thermosensitive mutants for which a rapid shift to a nonpermissive tem-

perature is used. Analysis of our data shows that the full extent of the effects of spliceosome mutations can be revealed only when the NMD (or any other RNA degradation system that also affects unspliced RNAs) is inactivated. Some transcripts such as *RPL16A* seem to be affected by all mutants whenever these are combined with NMD inactivation. However, some other transcripts such as *TUB1* seem to be affected by only a subset of mutants. Thus, splicing factor mutations seem to affect the splicing of genes to different extents, even when taking into account the effect of RNA degradation by NMD. We emphasize that our study may not have revealed the full extent of the phenotypes associated with the splicing mutants analyzed here because some other RNA degradation systems such as the nuclear exosome may also degrade some transcripts unaffected by NMD, as shown previously (Bousquet-Antonelli et al. 2000). Even when considering this possibility, our results underscore the importance of RNA surveillance in the quality control of splicing, as a large fraction of unspliced precursors generated by splicing defects are eliminated by RNA degradation.

MATERIALS AND METHODS

The *prp17Δ*, *prp18Δ*, *nam8Δ*, and *mud1Δ* deletion strains were obtained form Open Biosystems. The *prp22-1* mutant strain was from Vijayraghavan et al. (1989). Disruption of *UPF1* in splicing mutant background was performed using a PCR-based knockout strategy using the HIS marker (Longtine et al. 1998). Analysis of RNAs extracted from the single and double mutant strains by tiling arrays was performed as described (Sayani et al. 2008). Three independent biological replicates were obtained for each single or double mutant, except for the wild-type and *upf1Δ* strains for which four biological replicates were used (Sayani et al. 2008). Tiling array data are accessible in the GEO database (accession numbers GSE11621 and GSE18288). Tiling array signals were analyzed in MATLAB. Probe signals for the biological replicates for each strain were quantile normalized and averaged. The log2 of the ratio of the average of for each probe (perfectmatch— mismatch) was calculated for each strain comparison. Probes that hybridize 12-nt upstream of and 12-nt downstream from intronic regions were retained for intronic signal analysis. Z-scores were calculated in Matlab and are provided in Supplemental Table 1. Hierarchical clustering analysis using the average method was performed in JMP. Northern blot analysis was performed as described (Sayani et al. 2008). RT-PCR analysis was performed after DNase-I treatment of total RNA samples. A total of 40 μg of RNA was treated with 10 μL of Ambion RNase-free DNase I in a 100 μL reaction at 37°C for 1 h. The DNase-treated RNA samples were purified by phenol-chloroform extraction and precipitation and resuspended in 20 μL of water. Reverse transcription was performed using random hexamers. A total of 5 μg of DNase-treated RNA was incubated with 1 μL of random primers (50 ng/μL) in 12 μL. The reaction was heated to 65°C for 5 min and chilled on ice. Reverse transcription was performed using M-MLV Reverse Transcriptase in a 20 μL reaction for 50 min at 42°C. The reactions were incubated at 85°C for 5 min to heat inactivate the reverse-transcriptase and treated with RNase H for

20 min at 37°C. Two microliters of reverse transcription reactions were used for PCR amplification using standard procedure and including [α^{32}-P]dCTP for detection. Products were fractionated on polyacrylamide gels and the dried gels were exposed for phosphorimaging. The oligonucleotides used to amplify PCR products used for in vitro transcription and generation of riboprobes for Northern blots were the following:

RPL16A forward: GTCTGTTGAACCAGTTGTTGTC;
RPL16A reverse: GATCAATTAACCCTCACTAAAGGAGCAGTA
  GCATTAGCAGAGGCA;
RPS11B forward: AGGAAACGTATCAAGCCAGGGA;
RPS11B reverse: GATCAATTAACCCTCACTAAAGGTGGGGAG
  ACGTGGACTGGG;
RPS13 forward: ATGGGTCGTATGCACAGT; and
RPS13 reverse: GATCAATTAACCCTCACTAAAGGCCAGTTTG
  GTGGTAAGACAGCAAC.

The oligonucleotides for RT-PCR analysis were the following:

TUB1 forward: CGCCACCCAAGATCTGTAAACT;
TUB1 reverse: ACCAATCTGACAACCAGCTTGACC;
TAF14 forward: GTGACCTCGGAGCTGACTGATATT;
TAF14 reverse: TCTATACTCCATTGACGAACAGGA;
RPL16A forward: ACGGAAGGTGAAGAACAAATCGAG;
RPL16A reverse: AGCAACAACGGAAGCTAAACGACC;
RPS11B forward: CCACTGAATTAACTGTTCAATCTGAAAGAGC;
RPS11B reverse: GGTTCTCTTGGAAGTCTTGACCTTTGG;
ACT1 forward: AGAATAGGATCTTCTACTACATCAGCTT;
ACT1 reverse: ACGATAGATGGGAAGACAGCACGA;
RPS13 forward: AAATGGGTCGTATGCACAGT;
RPS13 reverse: AGGACAACTTGAACCAAGCTGGAG;
MER2 forward: GCTACTGGAACAAGATGCTGCTACGA;
MER2 reverse: CGTCGCCTTTCGATAACATTGCTG;
LBS3 forward: TCCAAGGAGTTTAAAGAGTGAGACA; and
LBS3 reverse: AACGTCTGGAGGAATGACTTGGTC.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at http://www.rnajournal.org.

ACKNOWLEDGMENTS

REFERENCES

Behm-Ansmant I, Kashima I, Rehwinkel J, Sauliere J, Wittkopp N, Izaurralde E. 2007. mRNA quality control: An ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett* **581:** 2845–2853.
Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R. 2008. Isolation of an active step I spliceosome and composition of its RNP core. *Nature* **452:** 846–850.
Bousquet-Antonelli C, Presutti C, Tollervey D. 2000. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102:** 765–775.
Burgess S, Couto JR, Guthrie C. 1990. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell* **60:** 705–717.
Chanfreau G, Legrain P, Dujon B, Jacquier A. 1994. Interaction between the first and last nucleotides of pre-mRNA introns is a determinant of 3′ splice site selection in *S. cerevisiae. Nucleic Acids Res* **22:** 1981–1987.
Chang YF, Imam JS, Wilkinson MF. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76:** 51–74.
Clark TA, Sugnet CW, Ares M Jr. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* **296:** 907–910.
Gottschalk A, Tang J, Puig O, Salgado J, Neubauer G, Colot HV, Mann M, Seraphin B, Rosbash M, Lührmann R, et al. 1998. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA* **4:** 374–393.
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* **90:** 7034–7038.
Hilleren PJ, Parker R. 2003. Cytoplasmic degradation of splice-defective pre-mRNAs and intermediates. *Mol Cell* **12:** 1453–1465.
Isken O, Maquat LE. 2008. The multiple lives of NMD factors: Balancing roles in gene and genome regulation. *Nat Rev Genet* **9:** 699–712.
Jaillon O, Bouhouche K, Gout JF, Aury JM, Noel B, Saudemont B, Nowacki M, Serrano V, Porcel BM, Segurens B, et al. 2008. Translational control of intron splicing in eukaryotes. *Nature* **451:** 359–362.
Jones MH, Frank DN, Guthrie C. 1995. Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. *Proc Natl Acad Sci* **92:** 9687–9691.
Konarska MM, Vilardell J, Query CC. 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol Cell* **21:** 543–553.
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.
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.
Liu L, Query CC, Konarska MM. 2007. Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nat Struct Mol Biol* **14:** 519–526.
Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae. Yeast* **14:** 953–961.
Mitrovich QM, Anderson P. 2000. Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in *C. elegans. Genes & Dev* **14:** 2173–2184.
Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M Jr. 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & Dev* **21:** 708–718.
Pikielny CW, Rosbash M. 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell* **41:** 119–126.
Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. 2007. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol* **5:** e90. doi: 10.1371/journal.pbio.0050090.
Rutz B, Seraphin B. 2000. A dual role for BBP/ScSF1 in nuclear pre-mRNA retention and splicing. *EMBO J* **19:** 1873–1886.
Sapra AK, Arava Y, Khandelia P, 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.

Sapra AK, Khandelia P, Vijayraghavan U. 2008. The splicing factor Prp17 interacts with the U2, U5 and U6 snRNPs and associates with the spliceosome pre- and post-catalysis. *Biochem J* **416:** 365–374.

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.

Scherrer FW Jr, Spingola M. 2006. A subset of Mer1p-dependent introns requires Bud13p for splicing activation and nuclear retention. *RNA* **12:** 1361–1372.

Tseng CK, Cheng SC. 2008. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* **320:** 1782–1784.

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.

Vijayraghavan U, Parker R, Tamm J, Iimura Y, Rossi J, Abelson J, Guthrie C. 1986. Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. *EMBO J* **5:** 1683–1695.

Vijayraghavan U, Company M, Abelson J. 1989. Isolation and characterization of pre-mRNA splicing mutants of *Saccharomyces cerevisiae*. *Genes & Dev* **3:** 1206–1216.

CHAPTER 3

Widespread and Regulated Use of Non-Productive Splice Sites

in *Saccharomyces cerevisiae*

Widespread and Regulated Use of Non-Productive Alternative Splice Sites in *S.cerevisiae*

Tadashi Kawashima[1], Stephen Douglass[2], Matteo Pellegrini[3] and Guillaume F. Chanfreau[1,4]

1: Department of Chemistry and Biochemistry and the Molecular Biology Institute, UCLA, Box 951569, Los Angeles CA 90095-1569

2: Bioinformatics Interdepartmental Program, UCLA, Box 951606, Los Angeles CA 90095-1606

3: Department of Molecular, Cellular and Developmental Biology, UCLA, Box 951606 Los Angeles, CA 90095-1606

4: To whom correspondence should be addressed

Abstract

The yeast *Saccharomyces cerevisiae* has been used as a model system to investigate the mechanism of pre-mRNA splicing but is thought to lack extensive use of alternative splice sites. Using RNA-Seq analysis of nonsense-mediated mRNA decay (NMD) mutants, we show many *S. cerevisiae* intron-containing genes exhibit alternative splice site usage, but most transcripts generated by splicing from these sites are non-functional, either because they introduce premature termination codons leading to degradation by NMD or because they encode non-functional proteins. Analysis of splicing mutants combined with NMD inactivation revealed the role of specific splicing factors in governing the use of these alternative splice sites, including a role for Prp18p in suppressing the use of a non-canonical 3'-splice site. Non-productive splicing can significantly reduce the overall splicing efficiency and expression of some transcripts, and the use of these sites can be increased during stress in a promoter-dependent manner to down-regulate these transcripts. These results reveal that the architecture of splicing signals in *S. cerevisiae* intron-containing genes is suboptimal, possibly to allow yeast cells to rapidly down-regulate gene expression in response to various environmental cues.

Introduction

Nonsense-mediated mRNA decay (NMD) is an RNA degradation system that degrades RNA containing premature termination codons (Isken and Maquat 2007; Kervestin and Jacobson 2012). 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 (Green, Lewis et al. 2003; Lareau, Inada et al. 2007; Ni, Grate et al. 2007). This interplay between alternative splicing and NMD is involved in the autoregulation of SR proteins (Green, Lewis et al. 2003; Lareau, Inada et al. 2007; Ni, Grate et al. 2007). 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 (He, Peltz et al. 1993; Mitrovich and Anderson 2000; Jaillon, Bouhouche et al. 2008; Sayani, Janis et al. 2008; Sayani and Chanfreau 2012). 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 basic components of the splicing machinery were identified through genetic screens in *S. cerevisiae* (Vijayraghavan, Company et al. 1989) and most splicing factors are highly conserved from yeast to mammalian cells (Wahl, Will et al. 2009). However similarities with mammalian splicing systems are limited by the perception that alternative splicing has a relatively minor impact on expression of spliced transcripts in *Ascomyceta*, and in particular in *S. cerevisiae*. Despite the presence of just under 300 intron-containing genes in *S. cerevisiae*, only a few examples of alternative splice (AS) site selection have been documented. The *SRC1* gene encodes an integral transmembrane protein, and 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 (Grund, Fischer et al. 2008; Mishra, Ammon et al. 2011). Alternative 3'-splice site selection has been shown to

regulate expression of the *APE2* gene according to temperature-dependent secondary structure of the transcript (Meyer, Plass et al. 2011). Recent work analyzing alternative splicing across different fungal species suggests 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 (Marshall, Montealegre et al. 2013). To the extent of our knowledge, the previously described functions of transcripts generated through alternative splice sites in *S. cerevisiae* have been restricted to the generation of different protein isoforms. In this study, we show that many *S. cerevisiae* intron-containing genes are spliced at non-productive splice sites, and that NMD degrades most mRNAs resulting from these splicing events. We show that the use of non-productive splicing events for some genes is upregulated during environmental stress, as a regulatory mechanism to down-regulate the levels of productively spliced transcripts from these genes in these specific conditions. These studies show that splice site selection in *S. cerevisiae* is much more flexible than previously thought, and that non-productive splice sites are widely generated and eliminated by RNA quality control mechanisms.

Results

RNA-Seq analysis reveals 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*, and that this is due to naturally with suboptimal splice sites in these genes (Sayani et al. 2008). To gain further insights into the function of NMD in proofreading splicing, 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 using BLAT(Kent

2002). This analysis revealed numerous occurrences of spliced transcripts arising from previously unknown splice sites, in both WT and the NMD mutants. For this study 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 canonical splicing events. RNA sequencing revealed that alternative splice sites were in general more abundant in the RNA samples obtained from the NMD mutant strains (Figure 1; Table S2). This is consistent with the fact that most of the 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). All three mutants showed an approximately 3.8-fold increase in unspliced intronic signal versus wildtype (Table S3), confirming our previous results from tiling arrays showing the involvement of NMD in eliminating unspliced transcripts genome-wide in *S. cerevisiae* (Sayani, Janis et al. 2008). 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 lower frequencies. The consensus sequences derived from the alternative splice sites identified in wild-type and all three mutants exhibited differences from the consensus sequences of the canonical splice sites (Fig1D,E). The alternative 5'-splice sites identified showed a relaxation of the conserved sequences, especially at positions 4 and 6 of the splice site compared to the consensus sequence obtained from the canonical 5'-splice sites. The 3'-splice sites 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 (Fig1D,E). Thus, the alternative splicing events identified in all four strains show decreased conservation in

their consensus sequences, suggesting that these might correspond to lower efficiency splice sites, and possibly to rarer events.

Strategy for alternative splicing events validation

The previous RNA-Seq analysis revealed the potential widespread usage of alternative splice sites (SS) that lead to PTC-containing transcripts. Figure 2 depicts specific mRNAs that were chosen for validation and further characterization. These transcripts were classified into three general categories: those with: 1) alternative 5' SS usage; 2) alternative 3' SS usage; and 3) a combination of both. Transcripts from category 1 included *RPL22B* as well as the previously reported *SRC1* (Grund, Fischer et al. 2008). Transcripts from category 2 included the genes encoding the RNA Polymerase III 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 category 3, 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 validated alternative splicing events by RT-PCR analysis using Cy3-end labeled PCR 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 performed with ^{32}P-end labeling with an appropriate ^{32}P-labelled ladder. This was necessary to confirm the sizes of all RT-PCR products and correlate the data back to gels obtained with Cy3 primers (data not shown). In addition to the wild-type and NMD-deficient *upf1Δ* strains, we chose to include other *S. cerevisiae* mutants that have been previously reported to affect splice site selection or splicing fidelity. Knockout mutants of genes encoding Mud1p and Nam8p were chosen for their association with the U1 snRNP and role in 5'-SS selection (Liao, Tang et al. 1993; Neubauer, Gottschalk et al. 1997;

Gottschalk, Tang et al. 1998; Puig, Gottschalk et al. 1999). The *HUB1* knockout was also included, as Hub1p was recently implicated in 5'-SS selection for *SRC1* (Mishra, Ammon et al. 2011). Prp17p and Prp18p were selected for their involvement in the 2nd step of splicing and their effects on 3'-SS selection (Umen and Guthrie 1995; Aronova, Bacíková et al. 2007). Finally, Isy1p was also included as a potential fidelity factor (Villa and Guthrie 2005).

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

As a first step in validating our RT-PCR strategy to assess the pattern of alternative splicing and the role of specific splicing factors, we first focused on *SRC1*, which exhibits two possible 5'-SS (Fig.2) and for which previous studies have demonstrated the role of various splicing factors (Grund, Fischer et al. 2008; Mishra, Ammon et al. 2011; Saha, Banerjee et al. 2012). RT-PCR analysis of *SRC1* splice variants confirmed the use of these two alternative 5'-SS (Fig. 3). Wild-type samples showed slightly more *SRC1-S* than *SRC1-L* transcript, consistent with previous reports (Grund, Fischer et al. 2008; Mishra, Ammon et al. 2011; Saha, Banerjee et al. 2012). Samples from the *upf1Δ* mutant showed a pattern similar to wild-type (Fig.3), indicating that both variants are stable and not NMD targets. 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 ratio of unspliced to spliced transcripts (Fig.3) due to reduced efficiency in the first step of splicing (Rodriguez-Navarro, Igual et al. 2002). The *prp17Δ* and *prp18Δ* mutants both showed decreased usage of *SRC1*-S 5' splice site, as suggested previously for the *prp17Δ* mutant at the protein level (Mishra, Ammon et al. 2011). 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 (Villa and Guthrie 2005). As reported previously (Mishra, Ammon et al.

2011; Saha, Banerjee et al. 2012), Hub1p inactivation resulted in reduced amounts of *SRC1-S*, coinciding with an increase in *SRC1-L* (Fig.3). Thus, the results described above confirmed the previously described effects of various splicing mutants on *SRC1* alternative splicing pattern and demonstrated 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, as well as unspliced *RPL22B* pre-mRNAs are present at higher levels in the *upf1Δ* mutant (Fig.3), indicating they are targeted by NMD. The large accumulation of unspliced species in the *upf1Δ* mutant indicates that the recognition of this splicing substrate might be inefficient, which may be the result of both the normal (GUACGU) and alternative (GUUUGU) 5'-SS having non-consensus sequences. Interestingly, the abundance of the alternative spliced product was found to decrease in double mutants with deletions of either of the U1 snRNP components Nam8p or Mud1p in the context of the *upf1Δ* deletion. The deletion of either one of these two factors might hinder the stability of the U1 snRNP binding to the alternative suboptimal 5' GUUUGU splice sites of *RPL22B*, resulting in a decreased usage. This is consistent with the known role of Mud1p and Namp1 in the first step of splicing (Puig, Gottschalk et al. 1999), and suggest their direct involvement in modulating 5'-SS selection of *RPL22B*. By contrast, no major splicing profile changes were observed in the *prp17Δ*, *prp18Δ*, *isy1Δ*, *hub1Δ* mutants, either alone or in combination with the *upf1Δ* deletion, showing the specificity of the effects detected with Nam8p and Mud1p. Thus, *RPL22B* exhibits two

competing sub-optimal 5'-SS, one of which is highly sensitive to perturbations in the U1 snRNP. The significance of this competition is further investigated below.

RT-PCR analysis shows antagonistic roles for Prp17p and Prp18p in the selection of non-productive alternative 3'-splice sites

TAN1 shows a UAG 3'-SS flanked on both ends by alternative 3' AAG sequences, both of which would generate PTC-containing transcripts marked for NMD degradation (Fig.2). The upstream AAG (AS 3' #1) is 6nt away from the normal 3' SS. The retention of 6nt 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 now generates an in-frame stop codon (Chanfreau 2010). The downstream AAG (AS 3' #2) is 7nt downstream of the normal 3'SS, resulting in a frame-shift 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 not detected unless NMD is inhibited (Fig.3). In samples from the *upf1Δ* strain, the signal of the two alternatively spliced products are similar in amount, and both species are detected at much lower levels than the normal splice product. Strikingly, the preference for these alternative 3' splice sites was altered when the Prp17p and Pprp18p factors were inactivated. Inactivating Prp17p resulted in increased use of the downstream AS 3' #2, while the upstream alternative splice site was no longer used. By contrast, Prp18p inactivation resulted in increased usage of the alternative 3'-SS most proximal to the branch point sequence (AS 3' #1), similarly to what has been observed by others for reporter transcripts (Zhang and Schwer 1997). 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 the pattern of usage of the two alternative 3'-splice sites of *TAN1*, showing the specificity of the effects observed with Prp17p and Prp18p.

Gapped sequence alignment showed that *TFC3* exhibits an alternative CAG 3'-SS 17nt downstream of the annotated AAG (Fig.2). This product can be detected in samples from the wild-type and splicing mutants, but is more abundant in the context of the *UPF1* deletion, showing that a large fraction of this product is degraded by NMD (Fig.3). When compared to transcripts spliced at the normal 3'-SS, this non-productive isoform represents a significant fraction of all spliced products, showing that a large fraction of the splicing events generate NMD-targeted, non-productive transcripts. We observed an increased accumulation of the downstream alternative 3'-splice product in the *prp17Δupf1Δ* mutant while inactivation of Prp18p had no effect on the splicing pattern, showing that this transcript is mostly dependent on Prp17p for proper 3'-SS selection. As expected, inactivation of the first step splicing factors Mud1p or Nam8p had no effect on the pattern of 3'-SS selected.

Analysis of the complex 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 5' GUACGU splice site results in the deletion of 30 nucleotides from the 3' end of exon 1 (Fig.2), 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 acid at positions 187-197, which occur within the most highly conserved region of this protein (Yan, Westfall et al. 2001). 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). By contrast, the other alternative transcript generated from splicing at a downstream AS 3' CAG

results in a potential PTC. However, this PTC-containing transcript would exhibit a short 85 bp 3'-UTR, which might render it insensitive to NMD to as suggested by the *faux 3' UTR* model (Muhlrad and Parker 1999; Amrani, Ganesan et al. 2004). Indeed, the abundance of this transcript was not increased in the *upf1Δ* mutant (Fig.3). However this transcript is also 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 (Yan, Westfall et al. 2001). 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* of *TFC3*.

GCR1 has the most complex splicing pattern of the selected targets. Gapped alignments identified an intronic GUAUGG alternative 5'-SS as well as an upstream CAG alternative 3'-SS (Fig.2). In addition to the 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, as well as a non-canonical AUG alternative 3'-SS 23nt further 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. The fact that these two 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, however was a very minor product (labeled as S$^{(annot.)}$ in Fig. 2 and 3). The major spliced product observed resulted from the use of the most upstream GUAUGG 5'-SS and 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 and this change is not expected to affect the function of the protein, as all *GCR1* mutations with phenotypic effects have been mapped to sites downstream of this short stretch of variation (Clifton and Fraenkel 1981; Holland, Yokoi et al. 1987; Uemura and Jigami 1995). Based on the abundance of S relative to S[(annot.)], we believe that S, and not S[(annot.)] is the main spliced product for the *GCR1* gene. It is surprising that a splice variant not detected in our RNA-Seq analysis appeared to be the major splice product. However, this can be explained by the limited sequencing coverage for this gene because this transcript is not highly expressed.

In addition to these major spliced products that are not subject to NMD, we also detected a series of alternatively spliced products degraded by NMD (as denoted by the asterisk 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 *upf1Δ* single and double mutant strains, and it is the most abundant of all *GCR1* alternatively spliced products subject to NMD (Fig.3). Another product is generated from combining the upstream GUAUGG 5'-SS with the most downstream UAG 3' splice site (*C in Fig.2). This product results in a PTC, as it introduces a translational frameshift which is not detected until the 43rd amino acid is translated. This product accumulates at low abundance in all samples and appears to be targeted by NMD, as its abundance increases slightly in *upf1Δ* samples, especially in the *prp17Δupf1Δ* strain, showing the importance of Prp17p in repressing this non-functional site.

Another set of NMD targets results from using the two most upstream 5'-SS combined with a highly unusual alternative AUG 3'-SS located further 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 confirmed through sequencing and RT-PCR

analysis of RNAs derived from *prp18Δupf1Δ* samples. First, 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. Second, 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 (Fig.S1). 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). Based on these results, we can unambiguously conclude that this AUG sequence is used as an alternative 3'-SS in the absence of Prp18p in combination with the first two 5' SS of *GCR1*. These results reveal the importance of Prp18p in ensuring proper 3'-SS selection for *GCR1* and repressing the use of non-canonical 3'-SS.

One final 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 *SRC1*, Hubp1p has no influence on their selection (Fig.3).

The use of the alternative 5'-splice sites of *RPL22B* and *GCR1* is increased in stress conditions.

The previous results validated our prediction that transcripts that are generated from alternative, non-productive splice sites are degraded by the NMD quality control system, and revealed the role of specific splicing factors in governing the choice among these alternative

sites. Interestingly, the sequence of some these non-productive splice sites was conserved across closely related yeast species (Fig.S2, *RPL22B* and Fig.S3, *TAN1*), suggesting that this sequence conservation might reflect a functional importance. Thus, we asked if the use of some of these alternative splice sites might be favored under certain conditions as a means to regulate gene expression. To test this, we monitored changes in splicing patterns for *RPL22B*, *TAN1*, and *TFC3* under diverse 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 (Pleiss, Whitworth et al. 2007; Bergkessel, Whitworth et al. 2011). In addition, many 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 to respond to the current stress condition (Li, Nierras et al. 1999; Gasch, Spellman et al. 2000). After 10 minutes of amino acid depletion, *RPL22B* WT (Fig.4A, lane 3) showed an increase in unspliced species as well as well as detectable levels of the alternatively spliced product in a wild-type context when compared to the SDC minimal medium control (Fig.4A, lane1). Upon NMD inactivation, the levels of unspliced as well as alternatively spliced products increased relative to the properly spliced, as would be expected when NMD transcripts are no longer degraded (Fig.4A lanes 2 and 4). By contrast, amino acid starvation had no effect on the splicing patterns of *TAN1* and *TFC3* and did not result in an increase in the levels of alternatively spliced species subject to NMD (Fig.4A). This result shows that the increase in the amount of alternatively spliced *RPL22B* transcript is not due to a decrease in NMD efficiency in these conditions, but rather to a switch in splice site selection.

We also investigated the effect of a 42°C heat shock (H.S.) on these splicing patterns. Under these conditions, *RPL22B* WT showed an increase in unspliced as well as a decrease in

the relative amount of spliced product (Fig.4A lane 5 vs. 7). More importantly the NMD defective strain *upf1Δ* showed an even larger increase in unspliced as well a large accumulation of alternatively spliced product that coincides with a decreased amount of canonical spliced product (Fig.4A lane 6 vs. 8). This result shows that in heat shock conditions, the amount of properly spliced is decreased while the use of the alternative splice site is being favored. By contrast, *TFC3* and *TAN1* exhibited an accumulation of unspliced species, but decreased levels of both the canonical and alternatively spliced species, consistent with a general inhibition of pre-mRNA splicing under heat shock (Yost and Lindquist 1991; Vogel, Parsell et al. 1995). Thus, the accumulation of the alternatively spliced *RPL22B* transcript under heat shock conditions described above is not due to a general stabilization of spliced forms subject to NMD, since the alternatively spliced isoforms of *TAN1* and *TFC3* that are normally degraded by NMD do not accumulate under these conditions.

Like heat shock, rapamycin treatment was shown to result in an inhibition of ribosomal protein gene splicing based on microarrays experiments (Bergkessel, Whitworth et al. 2011). Within 20 minutes of rapamycin treatment, *RPL22B* indeed showed similar trends as observed in heat shock, but to a much lesser degree, with an increase of unspliced species and of alternatively spliced *RPL22B* species, but no effect on the alternatively spliced *TAN1* and *TFC3* transcripts. Hyperosmotic shock (300mM LiCl exposure for 10min) only resulted in minimal effects. There were no changes observed for *TFC3* and *TAN1* targets under these stress conditions, and *RPL22B* showed only a slight increase in unspliced but the levels of spliced transcripts remained similar. Thus, *RPL22B* exhibits some change in the 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 pattern, especially in the absence of Prp18p, and because heat shock conditions resulted in the most dramatic changes in splicing pattern 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 even larger increase of the use of the non-canonical AUG site (*D species) and this product now constituted a large fraction of all spliced species. Thus, we conclude that *GCR1*, like *RPL22B* exhibits a switch in 5'-SS during heat shock, and that Prp18p function is essential to prevent the use of this non-canonical AUG site, especially during stress conditions.

The alternative 5'-splice site limits productive splicing of *RPL22B* in normal and stress conditions

Since *RPL22B* showed a splice site shift in favor of the alternatively spliced product over the normal splice site under heat shock conditions, we investigated the effect of mutations of the alternative 5'-SS, which is strongly conserved among yeast species (Fig.S2). We changed this suboptimal GUUUGU site to the consensus GUAUGU sequence at the endogenous chromosomal locus, and also introduced a chromosomal deletion of this splice site. Changing the alternative 5'-SS to the consensus GUAUGU sequence resulted in detectable amounts of alternative products at 25°C (Fig.5, lane 1 vs. 3.). Inactivation of NMD in this context showed that the major spliced product was now being generated from the use of the alternative consensus site, and splicing efficiency was increased to a point where unspliced species were no longer detected (Fig.5 lane 5). This result shows that the architecture of the two competing suboptimal 5'-SS naturally decreases the splicing efficiency of *RPL22B*, and that increasing the strength of the alternative 5'-SS is sufficient to enhance the overall splicing efficiency of this transcript. Under heat shock and NMD inactivation, this effect was even more prominent,

resulting in the alternatively spliced product being the only detectable species (Fig.5, lane 11). This result indicates that under heat shock conditions, *RPL22B* transcripts bearing the consensus alternative splice site mutation are now exclusively spliced 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'-splice site has been eliminated (Fig.5 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. In addition to RT-PCR, the same strains were analyzed by primer extension, which yielded a more quantitative view of the amount of transcripts generated under these conditions. Similar results were observed as those shown by RT-PCR, with the caveat that heat shock conditions resulted in much weaker signal than the samples at 25°C. This is due to the fact that under heat shock, many RPG mRNAs are downregulated (Li, Nierras et al. 1999; Pleiss, Whitworth et al. 2007; Bergkessel, Whitworth et al. 2011). However changing the alternative 5'-SS to a consensus sequence in the context of NMD inactivation was sufficient to recover a larger amount of spliced transcripts (Fig.5, lane 11, lower panel). Thus, alternative splicing, low splicing efficiency due to suboptimal 5' splice sites and NMD degradation appear to contribute to the general decrease in *RPL22B* levels as a means to rapidly halt production of this ribosomal protein, in the context of the overall decrease in ribosome biogenesis in order to shift energy and resources towards stress response necessary for survival.

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

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 the decrease in the use of the alternative 5'-SS (Fig.6 RT-PCR, lanes 3 and 4 vs. 5 and 6). This was particularly striking when considering the ratio between the canonical and alternatively spliced forms in the *upf1Δ* background under normal (Fig.6, lane 4 vs. 6) and heat shock (Fig.6, lane 12 vs. 10) conditions. Thus, alternative splicing regulation of *RPL22B* upon heat shock is mechanistically linked to the identity of the *RPL22B* promoter. The mechanism behind this promoter-dependent shift in alternative splice site selection is unclear. However, we can conclude from these results that transcriptional down-regulation and the increase 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 in *S. cerevisiae* results in 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 unexpectedly more complex than previously thought. 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*). Our observations imply that the rules of splice selection are intrinsically more flexible than previously thought, as the alternative splicing events that we have detected show significant deviations from the consensus

sequences (Fig.1). This is further illustrated by the finding that non-canonical 3'-SS such as an AUG site can be used (Fig.3) in the absence of Prp18p. 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* and *TAN1*, these transcripts represent a significant fraction of all transcripts generated from these loci. 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 significantly increase the level of splicing for this gene. Thus, the presence alternative and sometimes sub-optimal splice sites that compete with the normal splice sites contributes to an overall decrease in the amount of productively spliced transcripts. Because the overlap of the alternative splicing events detected between NMD-deficient strains was limited (Fig.1C), and because we detected by RT-PCR some alternative splicing events that escaped detection by RNA-Seq, we believe that we have not exhaustively identified the ensemble of splice sites that can be used by S. cerevisiae, and that many alternative splice sites can 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 allowed us to reveal some important and unexpected functions on the role of these factors on alternative splice site selection. We found that the Nam8p and Mud1p components are important contributors to improve 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 possess a suboptimal splicing signal, and therefore exhibits a weaker affinity for U1 binding and stronger requirement for Mud1p and Nam8p that impacts the efficiency of U1 snRNP assembly on the alternative splice site. Similarly, Prp17p and Prp18p were found to

have antagonistic roles on the use of alternative 3'-splice sites such as the ones exhibited by *TAN1*. Strikingly, we found that the absence of Prp18p resulted in the selection of a non-canonical AUG 3'-SS, 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 non-canonical 3'-SS. While we observed this function for *GCR1* only, we anticipate that a full genomic analysis of 3'-SS usage in the absence of Prp18p will reveal further examples of non-canonical 3' SS being used. Thus, our study reveals an important function for Prp18p in repressing the use of non-canonical splice sites and limiting aberrant splicing at non-YAG 3'-SS.

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 yeast species (Fig.S2 and S3), which, given the low conservation of intronic sequences in general, argues that this might reflects some degree of functional relevance. In addition, 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 contributes to the down-regulation of the normal *RPL22B* spliced transcript in stress conditions. The phylogenetically conserved, alternative, non-productive 5'-SS of *RPL22B* is functionally important not only because it reduces the overall splicing efficiency of *RPL22B*, but also because it participates in the down-regulation of *RPL22B* during stress. The transcriptional down-regulation of ribosomal proteins during stress has been documented previously (Li, Nierras et al. 1999). We show here that the promoter of the *RPL22B* gene is essential not only because it controls the

transcriptional repression during stress, but also because it drives the switch in 5'-SS selection to contribute to the overall repression of *RPL22B* during heat-shock. Thus, a combination of transcriptional and post-transcriptional regulations, through splicing inhibition (Pleiss, Whitworth et al. 2007; Bergkessel, Whitworth et al. 2011) and increased use of non-productive splice sites (this study) contributes to repress 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 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 much more complex than initially anticipated and contribute to genetic regulations and adaptations to environmental changes.

Materials and Methods

Yeast culture and RNA Analysis.

Yeast strains were grown at 25°C in YPD medium, unless indicated otherwise in the figures. Sample preparation and RNA sequencing was performed by Illumina. RT-PCR analysis and northern blot was performed as described (Kawashima, Pellegrini et al. 2009).

Mapping reads

All Solexa 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 (Kent 2002) 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 wildtype 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. 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 Excel (Microsoft Corp.).

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.

References

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. 2004. A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature 432(7013): 112-118.

Aronova, A., Bacíková, D., Crotti, L.B., Horowitz, D.S., and Schwer, B. 2007. Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing. RNA 13(9): 1437-1444.

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(8): 1461-1478.

Chanfreau, G.F. 2010. A dual role for RNA splicing signals. EMBO Rep 11(10): 720-721.

Clifton, D. and Fraenkel, D.G. 1981. The gcr (glycolysis regulation) mutation of Saccharomyces cerevisiae. J Biol Chem 256(24): 13074-13078.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11(12): 4241-4257.

Gottschalk, A., Tang, J., Puig, O., Salgado, J., Neubauer, G., Colot, H.V., Mann, M., Seraphin, B., Rosbash, M., Luhrmann, R., and Fabrizio, P. 1998. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. RNA 4(4): 374-393.

Green, R.E., Lewis, B.P., Hillman, R.T., Blanchette, M., Lareau, L.F., Garnett, A.T., Rio, D.C., and Brenner, S.E. 2003. Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics 19 Suppl 1: i118-121.

Grund, S.E., Fischer, T., Cabal, G.G., Antunez, O., Perez-Ortin, J.E., and Hurt, E. 2008. The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. J Cell Biol 182(5): 897-910.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. 1993. Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proc Natl Acad Sci U S A 90(15): 7034-7038.

Holland, M.J., Yokoi, T., Holland, J.P., Myambo, K., and Innis, M.A. 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(2): 813-820.

Isken, O. and Maquat, L.E. 2007. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev 21(15): 1833-1856.

Jaillon, O., Bouhouche, K., Gout, J.F., Aury, J.M., Noel, B., Saudemont, B., Nowacki, M., Serrano, V., Porcel, B.M., Segurens, B., Le Mouel, A., Lepere, G., Schachter, V., Betermier, M., Cohen, J., Wincker, P., Sperling, L., Duret, L., and Meyer, E. 2008. Translational control of intron splicing in eukaryotes. Nature 451(7176): 359-362.

Kawashima, T., Pellegrini, M., and Chanfreau, G.F. 2009. Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. RNA 15(12): 2236-2247.

Kent, W.J. 2002. BLAT--the BLAST-like alignment tool. Genome Res 12(4): 656-664.

Kervestin, S. and Jacobson, A. 2012. NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol 13(11): 700-712.

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446(7138): 926-929.

Li, B., Nierras, C.R., and Warner, J.R. 1999. Transcriptional elements involved in the repression of ribosomal protein synthesis. Mol Cell Biol 19(8): 5393-5404.

Liao, X.C., Tang, J., and 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(3): 419-428.

Marshall, A.N., Montealegre, M.C., Jimenez-Lopez, C., Lorenz, M.C., and van Hoof, A. 2013. Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. PLoS Genet 9(3): e1003376.

Meyer, M., Plass, M., Perez-Valle, J., Eyras, E., and Vilardell, J. 2011. Deciphering 3'ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. Mol Cell 43(6): 1033-1039.

Mishra, S.K., Ammon, T., Popowicz, G.M., Krajewski, M., Nagel, R.J., Ares, M., Jr., Holak, T.A., and Jentsch, S. 2011. Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature 474(7350): 173-178.

Mitrovich, Q.M. and Anderson, P. 2000. Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes Dev 14(17): 2173-2184.

Muhlrad, D. and Parker, R. 1999. Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. RNA 5(10): 1299-1307.

Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., Luhrmann, R., and Mann, M. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A 94(2): 385-390.

Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., and Ares, M., Jr. 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21(6): 708-718.

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(6): 928-937.

Puig, O., Gottschalk, A., Fabrizio, P., and Seraphin, B. 1999. Interaction of the U1 snRNP with nonconserved intronic sequences affects 5' splice site selection. Genes Dev 13(5): 569-580.

Rodriguez-Navarro, S., Igual, J.C., and Perez-Ortin, J.E. 2002. SRC1: an intron-containing yeast gene involved in sister chromatid segregation. Yeast 19(1): 43-54.

Saha, D., Banerjee, S., Bashir, S., and Vijayraghavan, U. 2012. Context dependent splicing functions of Bud31/Ycr063w define its role in budding and cell cycle progression. Biochem Biophys Res Commun 424(3): 579-585.

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(8): 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(3): 360-370.

Uemura, H. and Jigami, Y. 1995. Mutations in GCR1, a transcriptional activator of Saccharomyces cerevisiae glycolytic genes, function as suppressors of gcr2 mutations. Genetics 139(2): 511-521.

Umen, J.G. and 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(6): 584-597.

Vijayraghavan, U., Company, M., and Abelson, J. 1989. Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae. Genes Dev 3(8): 1206-1216.

Villa, T. and 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(16): 1894-1904.

Vogel, J.L., Parsell, D.A., and Lindquist, S. 1995. Heat-shock proteins Hsp104 and Hsp70 reactivate mRNA splicing after heat inactivation. Curr Biol 5(3): 306-317.

Wahl, M.C., Will, C.L., and Luhrmann, R. 2009. The spliceosome: design principles of a dynamic RNP machine. Cell 136(4): 701-718.

Yan, B.C., Westfall, B.A., and 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(15): 1383-1389.

Yost, H.J. and Lindquist, S. 1991. Heat shock proteins affect RNA processing during the heat shock response of Saccharomyces cerevisiae. Mol Cell Biol 11(2): 1062-1068.

Zhang, X. and Schwer, B. 1997. Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. Nucleic Acids Res 25(11): 2146-2152.

Figure legends

Figure 3.1

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. Sequence logo analysis of 5'- and 3'- splice sites for all canonical splicing events detected by RNA-Seq in wild-type and NMD mutant strains.

E. Sequence logo analysis of 5'- and 3'- splice sites for all non-canonical splicing events detected by RNA-Seq in wild-type and NMD mutant strains.

Figure 3.2

Spliced isoforms 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 5' SS of *SRC1* is located 4nt upstream from the annotated 5' SS. The alternative 5' SS of *RPL22B* is located 64nt downstream from the annotated 5' SS. The alternative 3' SS of *TAN1* are located 6nt upstream and 7nt downstream from the annotated one. The alternative 3' SS of *TFC3* is located 17nt downstream from the annotated 3' SS. The alternative 5' and 3' SS of *GPI15* are located 36nt downstream and 14nt upstream, respectively, from the annotated 5' and 3' SS. The alternative 5' SS of *GCR1* are located 5nt upstream (GUAUGG); 51nt

downstream (GUAUGG) and 627nt downstream from the annotated 5'SS. The alternative 3' SS of *GCR1* are located 40nt upstream (AUG) and 17nt downstream (CAG) from the annotated 3'SS.

Figure 3.3

RT-PCR analysis of alternatively spliced products produced from the *SRC1*, *RPL22B*, *TAN1*, *TFC3*, *GPI15* and *GCR1* genes 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.

Figure 3.4

RT-PCR analysis of alternatively spliced products under stress conditions

A. Analysis of species from the *RPL22B*, *TAN1*, and *TFC3* genes in various stress conditions.

B. Analysis of species from the *GCR1* gene in heat-shock conditions.

Figure 3.5

Effects of mutations of the *RPL22B* alternative 5' splice site on *RPL22B* splicing and expression under normal and heat shock conditions

N, natural 5'-splice site; C, consensus GUAUGU; Δ = deletion.

Figure 3.6

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 strains in glucose (YPD) or galactose (YPGal) containing media. N= natural *RPL22B* promoter; G = Galactose promoter.
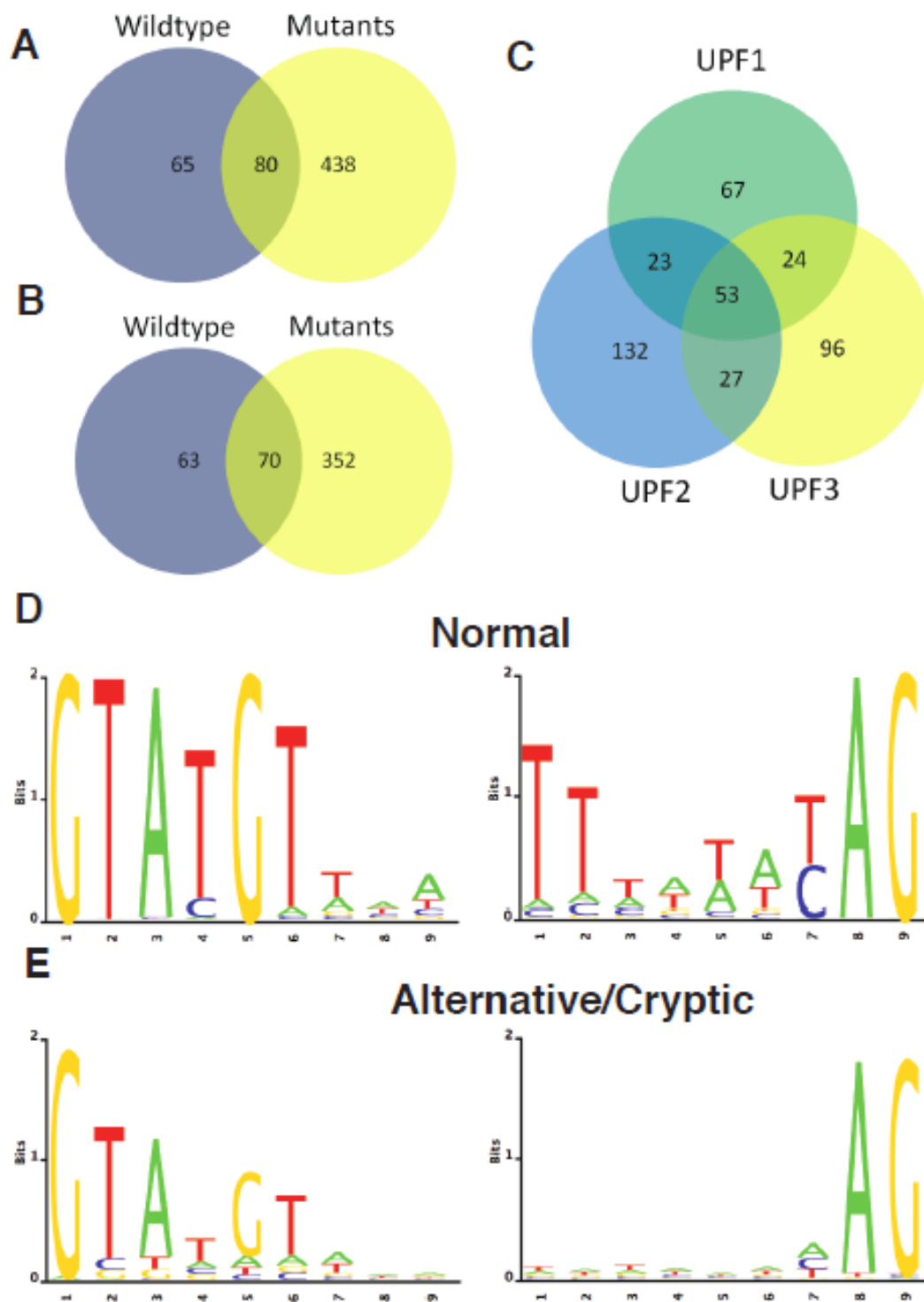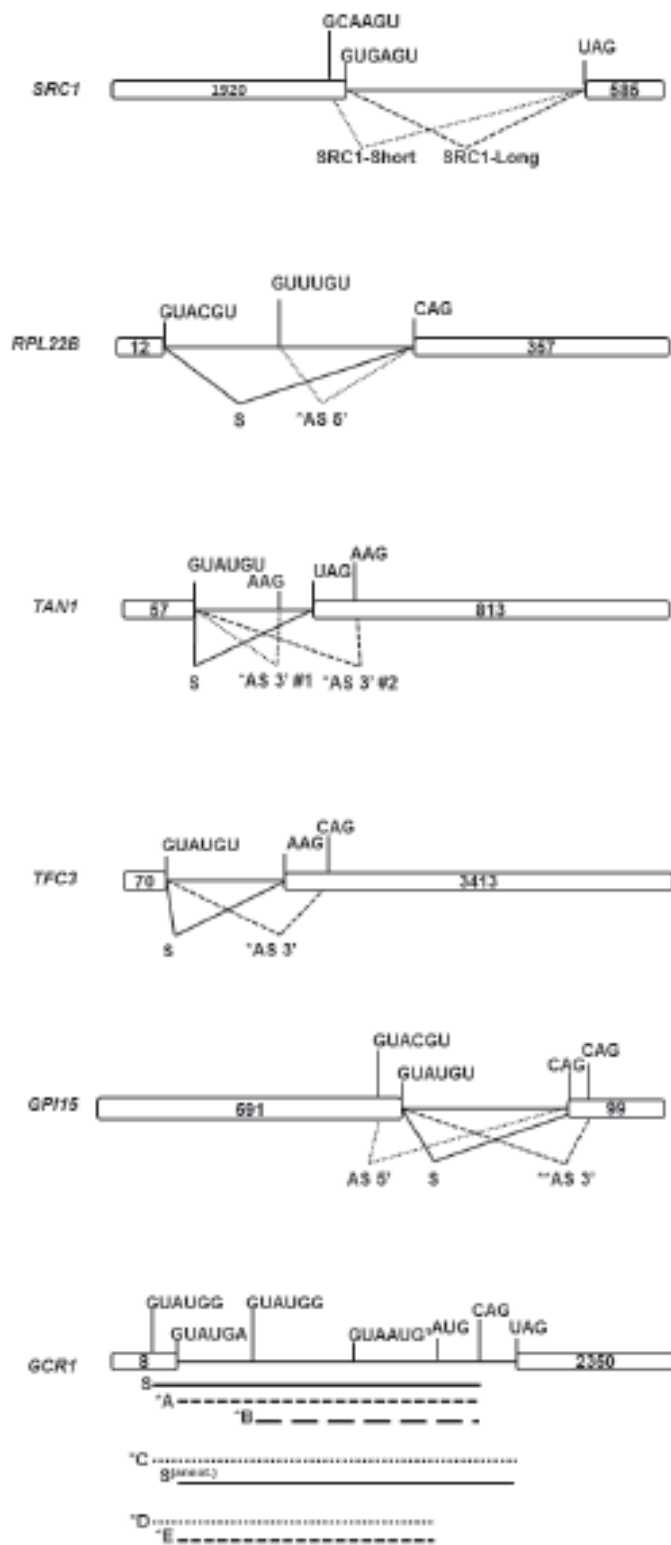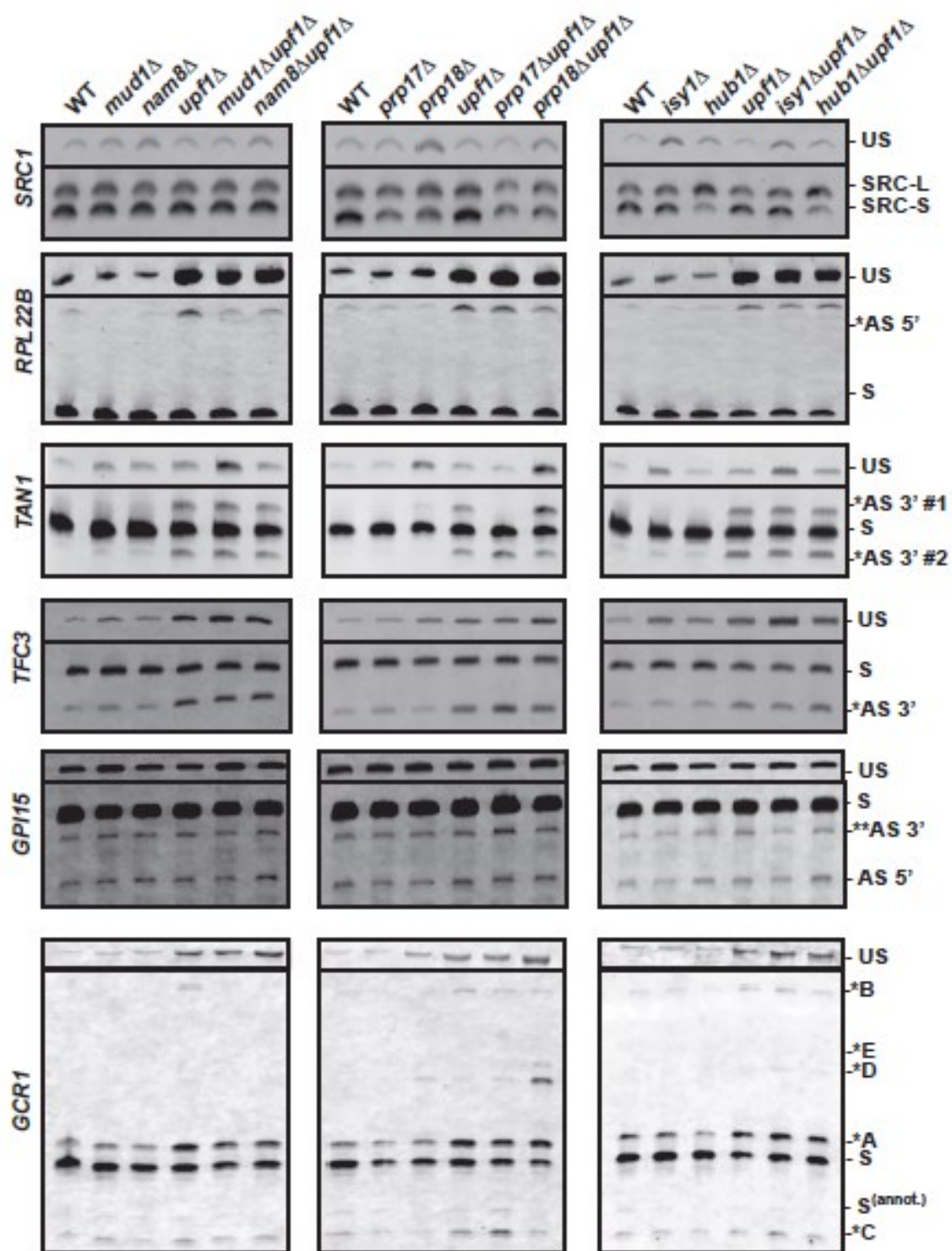
Figure 3.1

Figure 3.2

Figure 3.3

Figure 3.4

Figure 3.5

Figure 3.6

		25°C						42°C, 20min				
		YPD		YPGal				YPD		YPGal		
UPF1	N	Δ	N	Δ	N	Δ	N	Δ	N	Δ	N	Δ
Promoter	N	N	N	N	G	G	N	N	N	N	G	G

US — *RPL22B* RT-PCR

*AS 5'

S

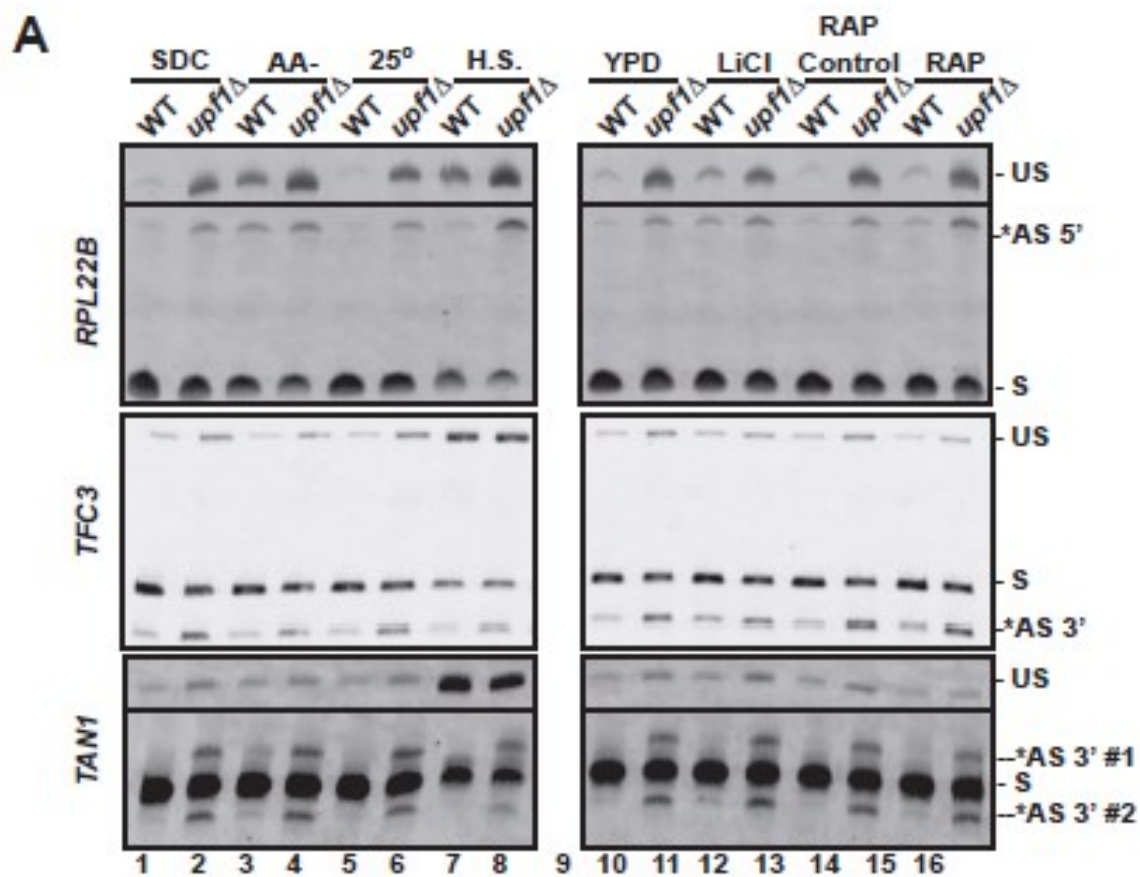Lanes: 1 2 3 4 5 6 7 8 9 10 11 12
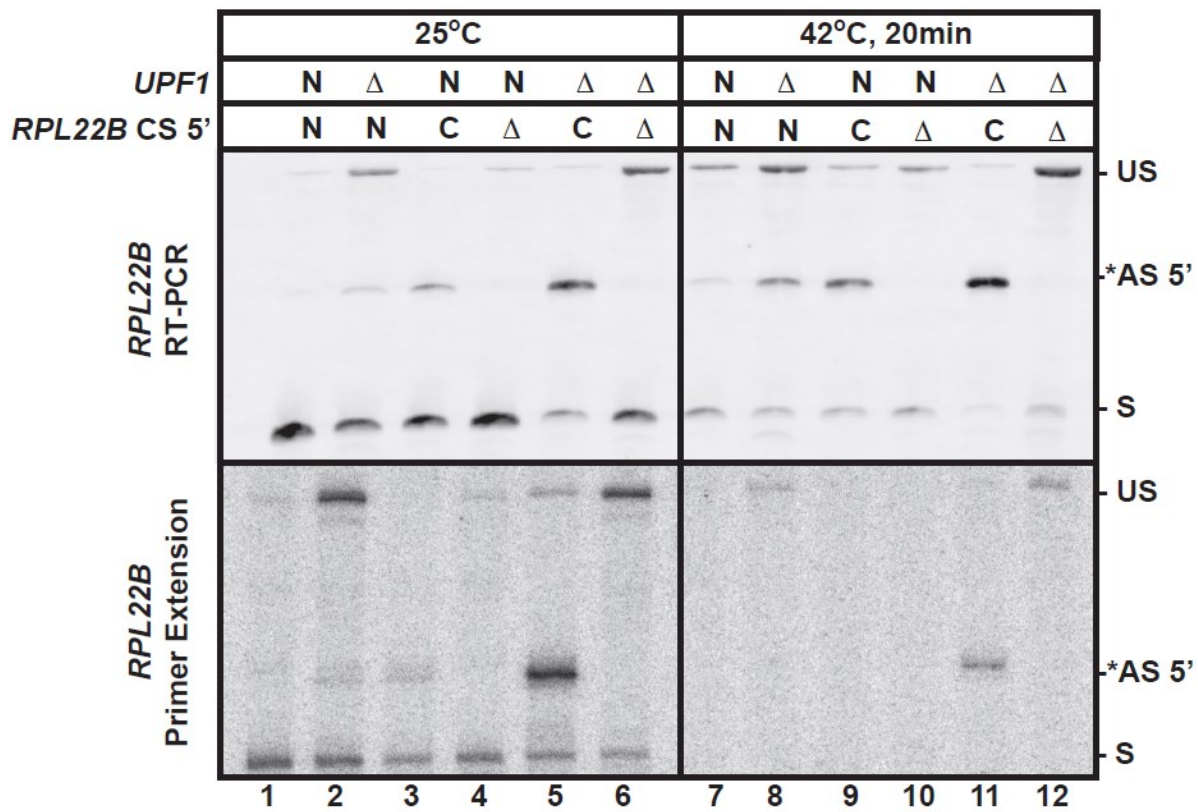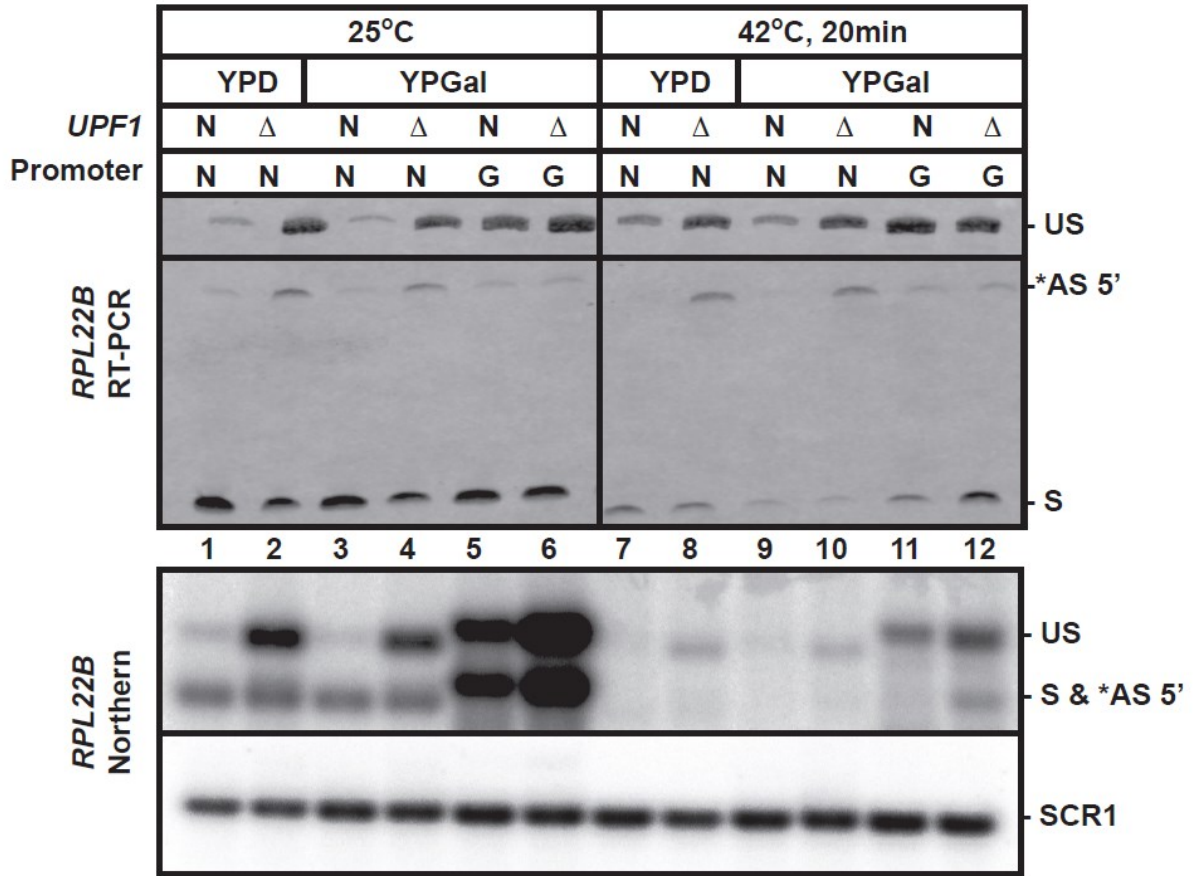
RPL22B Northern

US

S & *AS 5'

SCR1
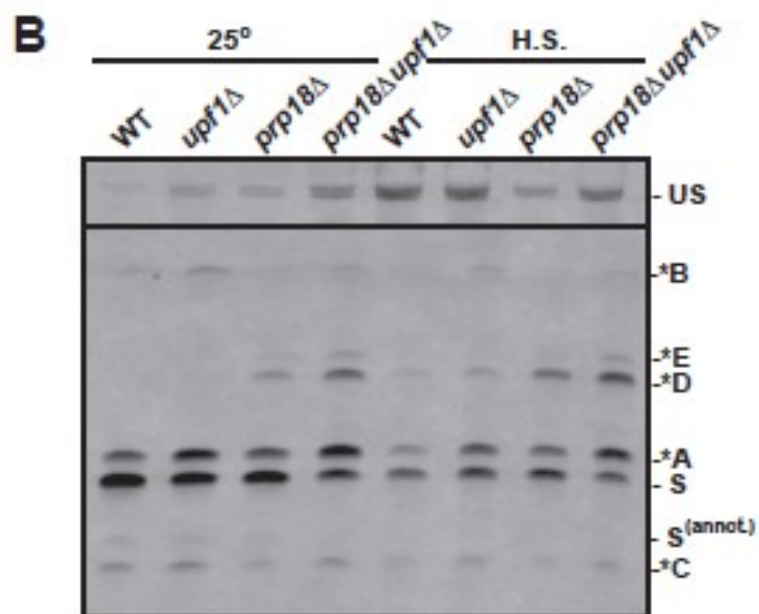
CHAPTER 4

Quality control of *MAT*a1 Splicing and Exon Skipping

by Nuclear RNA Degradation

Quality control of *MATa*1 splicing and exon skipping by nuclear RNA degradation

Defne E. Egecioglu, Tadashi R. Kawashima and Guillaume F. Chanfreau*

Department of Chemistry & Biochemistry and the Molecular Biology Institute, University of California, Los Angeles CA 90095-1569, USA

ABSTRACT

The *MATa*1 gene encodes a transcriptional repressor that is an important modulator of sex-specific gene expression in *Saccharomyces cerevisiae*. *MATa*1 contains two small introns, both of which need to be accurately excised for proper expression of a functional *MATa*1 product and to avoid production of aberrant forms of the repressor. Here, we show that unspliced and partially spliced forms of the *MATa*1 mRNA are degraded by the nuclear exonuclease Rat1p, the nuclear exosome and by the nuclear RNase III endonuclease Rnt1p to prevent undesired expression of non-functional a1 proteins. In addition, we show that mis-spliced forms of *MATa*1 in which the splicing machinery has skipped exon2 and generated exon1–exon3 products are degraded by the nuclear 5′–3′ exonuclease Rat1p and by the nuclear exosome. This function for Rat1p and the nuclear exosome in the degradation of exon-skipped products is also observed for three other genes that contain two introns (*DYN2*, *SUS1*, *YOS1*), identifying a novel nuclear quality control pathway for aberrantly spliced RNAs that have skipped exons.

INTRODUCTION

The protein product of the *MATa*1 gene in *Saccharomyces cerevisiae*, also called 'a1', is a transcriptional regulator, and, together with the alpha2 protein, specifically acts to repress the expression of haploid-specific genes as well as the HO gene in wild-type diploid cells (1). Both of these proteins are also normally expressed in their respective haploid cells, under the haploid-specific transcriptional program for either *MATa* or MATalpha cells

(Figure 1A). *MATa*1 has not been shown to have a specific function in haploid *MATa* cells (although it is expressed constitutively), but is thought to play a role in the expression of mating pheromones (2). On the other hand, alpha2 represses *MATa*-specific gene expression in MATα cells. In the case of diploid cells, these two homeodomain containing proteins heterodimerize and bind DNA cooperatively in order to repress the haploid-specific transcriptional program (3). *MATa*1 contains a homeodomain region between residues 66 and 126, which is important for a1 biological function. Structural studies have shown that within the a1 protein, Helices 1 and 2 of the homeodomain are important for maintaining contact with alpha2, whereas Helix 3 is important for mediating interactions with DNA (4). While a1 alone does not display any sequence specific DNA binding properties, the a1-alpha2 heterodimer has 3000-fold increase in sequence recognition when compared to alpha2 alone (5,6).

The gene encoding *MATa*1 is unusual, as it is one of the few *S. cerevisiae* genes containing two introns (7 9) (Figure 1A). In addition, these introns are smaller than most *S. cerevisiae* introns. Because exon3 is also small, previous studies have shown that partially spliced *MATa*1 mRNAs that have excised intron 1 but retain intron2 have the potential to encode a protein which is actually longer than the wild-type version and is inactive (8). Thus, it is important that the splicing of *MATa*1 occurs accurately, in order to prevent the accumulation of unspliced or partially spliced forms that would be non-functional. In this study, we show that unspliced and partially spliced forms of *MATa*1 are degraded by nuclear RNA turnover pathways. In addition, we demonstrate the existence of a novel RNA degradation mechanism that specifically targets exon2 skipped products of *MATa*1 and of three other genes that contain two introns (*DYN2*, *SUS1*, *YOS1*). These studies show that
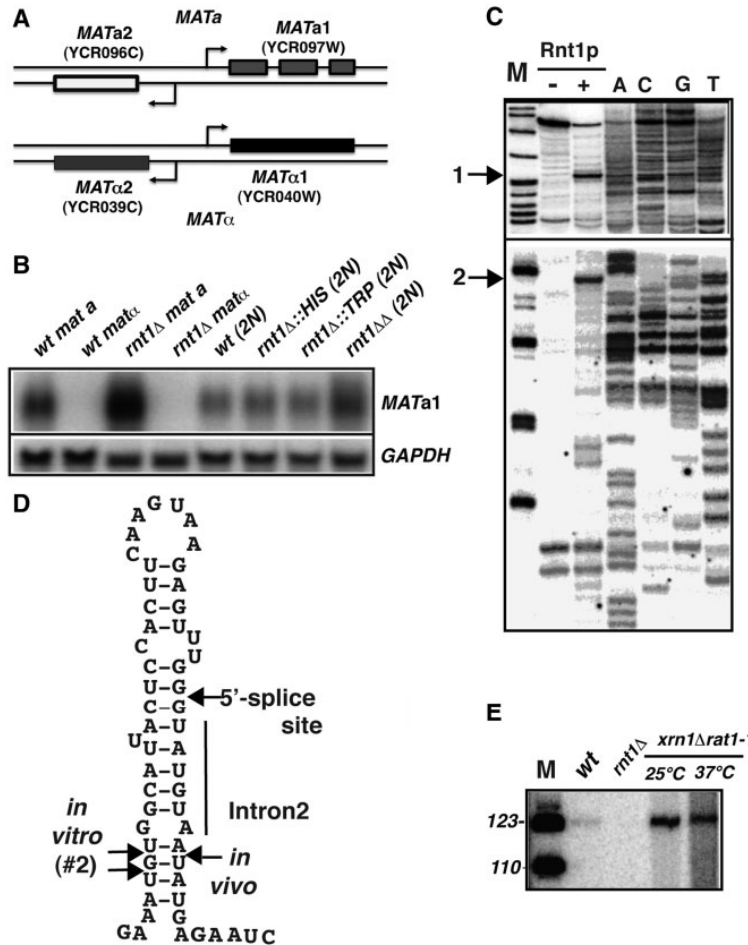
Figure 1. Architecture of the *MAT* locus, analysis of *MAT*a1 expression in Rnt1p mutant strains and cleavage by Rnt1p. (**A**) Schematic representation of gene structures at the *MAT*a and *MAT*alpha loci. (**B**) Northern blot analysis of *MAT*a1 expression in wild-type and *rnt1Δ* strains in the SK1 background. 2N indicate diploid cells. (**C**) *In vitro* cleavage of a model *MAT*a1 substrate by recombinant Rnt1p. Shown is a primer extension analysis of RNAs incubated in buffer or with recombinant Rnt1p. A sequencing ladder was obtained using the same oligonucleotide. 1 and 2 indicate the location of the major cleavage products. No other cleavage products were observed in the central portion of the gel which is not shown. (**D**) Predicted Rnt1p stem loop structure. Shown are the locations of the cleavage sites mapped *in vitro* (panel C) or *in vivo* (panel E). (**E**) RT PCR analysis of Rnt1p cleavage intermediates. RNAs presenting a 5′-phosphate group from the corresponding strains were ligated to an adaptor RNA, amplified using a *MAT*a1-specific reverse primer and an adaptor forward primer, and fractionated on an acrylamide gel.

the splicing of *MAT*a1 is relatively inefficient, and that nuclear degradation mechanisms have evolved to limit the accumulation of mRNAs that have escaped the splicing machinery or that have been aberrantly spliced.

MATERIALS AND METHODS

Yeast strains and media

Strains of mating type-a (*MAT*a) were utilized, unless otherwise indicated. Yeast strains were grown at 30°C until mid-log phase (OD600 = 0.4–0.6), in YPD medium unless indicated otherwise. The majority of the strains used in this study were in the BMA64 (10) or BY4741 (Open Biosystems) backgrounds and were described previously (11,12). The SK1 (13) derived strains used in this study were:

KBY257 (n) MATa, wild-type; KBY258 (n) MATα, wild-type; KBY257 (n) *rnt1::HIS*; KBY258 (n) *rnt1::TRP*; KBY (2n) *rnt1::HIS, RNT1*; KBY (2n) *rnt1::TRP, RNT1*; KBY (2n) *RNT1, RNT1*; KBY (2n) *rnt1::HIS rnt1::TRP*. SK1 strains were grown in 30°C at steady state. The *xrn1Δ::URA rat1-1* (MATa) strain was obtained from the Tollervey Lab (14), and was grown at 25°C until mid-log phase, and then shifted to 37°C for 3 h utilizing pre-warmed media. For generation of the *rnt1::KAN* replacement in the *xrn1Δ rat1-1* strain, PCR-based disruption of the *RNT1* gene was performed as described in Longtine *et al.* (15) through the amplification of the KAN[R] marker.

RNA analysis

The list of oligonucleotides used in this study is provided in Supplementary Table S1. Templates for *in vitro* RNA transcription were obtained through PCR amplification of *MATa1* sequences and incorporating a T3 RNA Polymerase promoter sequence at their 5'-ends. *In vitro* transcription was performed using the T3 MEGAScript or MAXIScript kits (Applied Biosystems/Ambion), and the resulting RNAs were purified through phenol::chloroform extraction, and checked on both native and denaturing gels prior to further experimentation. *In vitro* cleavage assays with purified recombinant His6-Rnt1p were carried out generally as described (16). Primer extension on cleaved *in vitro* transcribed RNAs for mapping cleavage sites was performed as described (16).

Total RNA extraction from yeast under denaturing conditions and subsequent northern blotting analysis were performed as described (16). Detection of RNA species on northern blots was performed using either 5'-labeled oligonucleotides, PCR probes, or riboprobes. Riboprobes were generated using the T3 MAXI kit (Applied Biosciences/Ambion) and hybridized to membranes as described (17).

Primer extension on total RNAs was performed as in (11,18) using 5 10 μg of denatured total RNA. The various primers used for reverse transcription and PCR amplification are listed in Supplementary Table S1. An amount of 40 μg of total RNA were digested with Ambion RNase-free Turbo DNase (8 U enzyme) for 45 min at 37°C in 200 μl. The reaction was phenol chloroform extracted with equal volume organic and ethanol-precipitated. The digested RNA was then resuspended in 20 μl water and quantitated. An amount of 5 μg of DNase treated RNA was combined with 0.4 μl (25 mM each) dNTPs and 1 μl of 50 ng/μl random primer mix to a total volume of 12 μl. The mixture was heated to 65°C for 5 min, followed by placing on ice for 2 min. This cooled mixture was combined with Invitrogen RNaseOUT (40 U) and Invitrogen M-MLV Reverse Transcriptase (200 U). RT reaction was left at 25°C for 10 min, followed by 42°C incubation for 50 min. The reaction was terminated by heating at 85°C for 5 min. PCR reactions included 1.5 μl of this reaction, 500 nM Cy3 forward and reverse primers, and were incubated at 95°C for 3 min, 35 cycles of (30 s at, 95°C, 30 s at 58°C, 30 s at 72°C), and a final incubation at 72°C for 3 min. Equal volumes of RT PCR products were combined with 95% formamide/dye mixture and denatured at 95°C for 10 min. This mixture was then loaded onto a 5% denaturing acrylamide gel (1× TBE). Gels were dried prior to scanning on a Molecular Imager FX System. Poly(A) enrichment of total RNAs was performed using the Ambion poly(A) purist MAG kits as described (12).

Amplification of cleavage products and sequencing analysis

For the adaptor ligation of cleavage products carrying 5'-phosphate groups, an adaptor RNA was purchased from Dharmacon. Total RNAs from various strains were ligated with the adaptor RNA on their 5'-phosphate ends, reverse transcribed using the E3 REV primer, and amplified by PCR. The primer pair for the PCR consisted of a primer that was the DNA version of the adaptor sequence, and the E3 Rev primer.

RESULTS

Accumulation of *MATa1* transcripts in *rnt1Δ* cells

Previous microarray analysis of the transcriptome of cells lacking the double-stranded RNA endonuclease Rnt1p revealed that Rnt1p triggers the degradation of mRNAs involved in iron uptake and homeostasis as well as two unspliced pre-mRNAs, *RPS22B* and *RPL18A* (16,19). To investigate if Rnt1p targets additional unspliced precursors for degradation, we mined these microarrays to search for additional intron-containing transcripts which would be subject to Rnt1p-mediated degradation. Such transcripts are more abundant in cells lacking Rnt1p activity because the absence of Rnt1p-mediated nuclear degradation of the unspliced pre-mRNA prevents competition with splicing, and increases the pool of RNA substrates directed towards the splicing pathway. These microarrays showed that the *MATa1* mRNA was expressed at higher levels in the *rnt1Δ* strain compared to wild-type. In order to confirm the microarray data, we analyzed *MATa1* expression in wild-type and *rnt1Δ* cells in both *MATa* and *MATalpha* mating types in the SK1 background strain. This strain is frequently used to study events of meiosis and sporulation due to the fast and effective sporulation properties of their diploids (13). Northern analysis showed that the *MATa1* mRNA accumulated at least four times more in *rnt1Δ* cells than in wild-type cells (Figure 1B). In the case of *MATalpha* cells, no *MATa1* signal was detected for either wild-type or *rnt1Δ* cells, which showed that the accumulation of *MATa1* in the *rnt1Δ* strain was not due to a derepression of the silent mating-type locus *HMRa-1*. The *MATa1* transcript was also three times more abundant in the diploid *rnt1ΔΔ* strain than in the isogenic wild-type diploid cells or heterozygous strains (Figure 1B). Agarose northern blots, such as the ones shown in Figure 1 were not able to separate unspliced precursors from spliced mRNAs, because of the small size of *MATa1* introns. As shown below, the accumulating *MATa1* in *rnt1Δ* cells seemed to result from at least two separate products, one being the mature mRNA, and the other possibly the unspliced, or a partially spliced versions of the transcript.

Rnt1p cleaves a non-canonical stem–loop structure that encompasses the 5'-splice site of intron2

The previous experiments suggested that Rnt1p might cleave the *MATa1* precursor transcript to eliminate unspliced species. To further test this hypothesis we generated an *in vitro* transcript containing the *MATa1* exons and introns, and incubated it with recombinant Rnt1p or with buffer alone. We found two major cleavage sites for Rnt1p in this model *MATa1* transcript (Figure 1C). The first one (#1, Figure 1C) did not match to any specific structure. However, the second major site

(#2, Figure 1C) corresponded to a sequence on a predicted stem loop structure, which bridges the second exon and the second intron and contains the 5′-splice site of the second intron (Figure 1D). The location of this cleavage site fits well with the properties of Rnt1p, which has been shown to cleave 14- to 16-bp away from terminal loops (20). However, the terminal loop found in this predicted structure is not a canonical AGNN tetraloop, which Rnt1p has been shown to bind and recognize (20,21). Instead, a 7-nt loop containing the AGNN motif is present (Figure 1D).

To detect a corresponding Rnt1p cleavage product *in vivo*, we used a linker-mediated RT PCR technique, which allows the detection of RNase III cleavage products by taking advantage of the 5′-phosphate groups generated by RNase III cleavage, a strategy similar to the RNA adapter ligation methods used in 5′-RACE and miRNA cloning kits (Ambion). These cleavage products are normally unstable because they are rapidly degraded but inactivation of the Xrn1p and Rat1p 5′ 3′ exonucleases allows their detection (14,16,18). Using this technique, we detected small amounts of a specific product in wild-type cells, the abundance of which was largely increased in an *xrn1Δrat1-1*strain (Figure 1E). This product was completely undetectable in samples extracted from a *rnt1Δ* strain, further suggesting that it corresponds to a Rnt1p cleavage product. Precise sizing analysis of this product showed that the adaptor was ligated to a sequence corresponding to the 3′-side of the predicted stem loop (Figure 1D; *in vivo* arrow). This sequence was found to be directly opposite to the site where the cleavage product was detected *in vitro* (Figure 1C). Thus, the *in vitro* cleavage and *in vivo* mapping experiments results fit well to demonstrate that Rnt1p cleaves a structure within the *MAT*a1 transcript that bridges the second exon and the second intron.

Accumulation of *MAT*a1 transcripts in nuclear RNA degradation mutants

Previous studies had shown that degradation of unspliced species of *RPS22B* relies on Rnt1p, but also on the action of the 5′–3′ exonucleases Xrn1p and Rat1p (16). To further investigate the mechanisms of degradation of *MAT*a1 transcripts, we analyzed its expression in a panel of RNA degradation mutants. This included the *rat1-1* strain, in which Rat1p is inactivated using a thermosensitive allele (14), the *xrn1Δ* strain, which carries a deletion of the gene encoding the cytoplasmic Xrn1p exonuclease, and the *rrp6Δ* strain containing a deletion of the nuclear exosome component Rrp6p. The *rnt1Δ* strain was included for comparison. We also analyzed *MAT*a1 levels in a combination of double or triple mutant strains (Figure 2). Northern analysis of *MAT*a1 expression in the *xrn1Δ rat1-1* strain showed an increase of *MAT*a1 transcripts at permissive temperature (Figure 2A), which was slightly lower than the accumulation observed in the *rnt1Δ* strain. However when this strain was combined with the *rnt1Δ* deletion, we observed a dramatic accumulation of signal in the amount of *MAT*a1 transcripts. Interestingly, we also

observed a large increase of *MAT*a1 signal in the *rnt1Δrrp6Δ* strain at 25°C, as well as in the triple mutant combining the *rrp6Δ* deletion to the *xrn1Δ rat1-1* strain. The RNA species detected for *MAT*a1 in the *xrn1Δ* strain migrate slightly faster than those detected in the other strains, possibly because they correspond to deadenylated forms.

When these strains were shifted to 37°C, we observed a very strong accumulation of *MAT*a1 transcripts in the *xrn1Δ rat1-1* strain, which was comparable to that observed in the *rnt1Δ* strain grown in the same conditions. We also observed an increase of signal in the *rnt1Δ* strain compared to 25°C. The combination of these mutants with other mutations did not result in an increase of signal, suggesting that Rnt1p and Rat1p play a major role in the degradation of *MAT*a1 transcripts when these strains are shifted to 37°C. Because of the small size of *MAT*a1 introns, the poor resolution of spliced and unspliced species in agarose northern blots prevented us from concluding whether the mature *MAT*a1 or unspliced species are the predominant species accumulating in RNA degradation mutants, or if an accumulation of both contribute to the large increase of signal. To investigate if unspliced *MAT*a1 transcripts contribute to the increase of signal observed in Figure 2A for the different mutant strains, we tried to hybridize membranes with probes specific to intronic sequences. This proved to be challenging, since the small size of each *MAT*a1 intron (54 and 52 nt) made the use of riboprobes or random primed probes impractical. In addition the low complexity of *MAT*a1 intronic sequences made it difficult to design oligonucleotide probes. However one sequence in intron2 allowed us to detect unspliced and/or partially spliced species of *MAT*a1 containing intron2 (Figure 2B). This probe detected the accumulation of unspliced species in the *rnt1Δ* strain at 25°C, which was further increased in the *xrn1Δ rat1-1rnt1Δ* triple mutant strain. When shifted to 37°C, we also observed a stronger intronic signal in the *rnt1Δ* strain, and in the *xrn1Δ rat1-1* strain and in its derivatives (Figure 2B). This result suggested that Rnt1p, as well as Rat1p contribute to degrading unspliced species of *MAT*a1. Based on the comparison of the patterns observed in Figure 2A and B, it is likely that the large increase in signal observed for the various degradation mutants is the result of an increase in both unspliced and spliced species of *MAT*a1.

Analysis of *MAT*a1 expression by RT–PCR confirms the accumulation of unspliced forms of *MAT*a1 in Rnt1p and Rat1p-deficient cells

The previous data suggested that cleavage of the *MAT*a1 pre-mRNA transcript by Rnt1p would initiate degradation of unspliced or partially spliced transcripts that still retain intron2, and also revealed a role for nuclear degradation by Rat1p and the nuclear exosome in eliminating these species. However, it was difficult to fully characterize the different RNAs accumulating in the various strains by northern blot because of their relatively low abundance, their poor resolution and the difficulty to use intron specific probes. To circumvent these
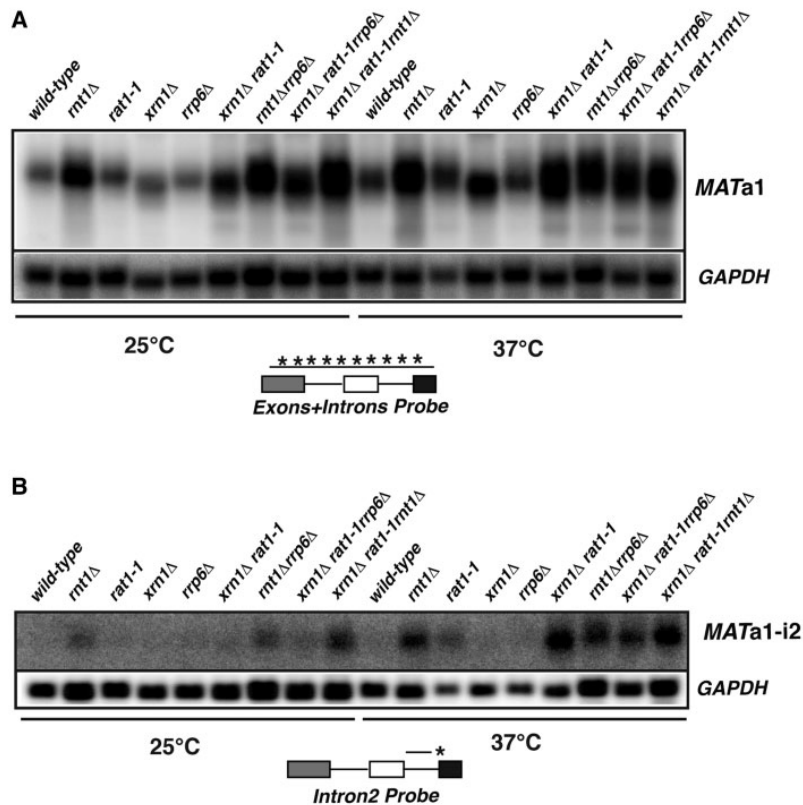
Figure 2. Northern blot analysis of *MAT*a1 expression in *rnt1Δ* and exonuclease mutant strains. The isozymes of the Glyceraldehyde 3 Phosphate Dehydrogenase enzyme (G3PDH, or GAPDH) were used as a loading control. (**A**) The membrane was hybridized with a riboprobe covering most of the *MAT*a1 sequence (3 exons and 2 introns) prior to GAPDH hybridization. (**B**) The membrane was hybridized with an intron2-specific oligonucleotide probe (Supplementary Table S1).

problems, we analyzed *MAT*a1 expression in all strains by RT–PCR (Figure 3A). This analysis confirmed the accumulation of fully unspliced and partially spliced transcripts still retaining intron 2 in the *rnt1Δ* strain (Figure 3A). We did not observe by RT–PCR the higher levels of spliced *MAT*a1 mRNAs in the *rnt1Δ* strain detected by northern blots in Figures 1 and 2. We hypothesized that the number of PCR cycles used to detect unspliced species led to a saturation of the signal for the spliced mRNA in these experiments. Using a smaller number of PCR cycles, we were able to confirm higher levels of spliced mRNAs in the *rnt1Δ* strain compared to the wild-type (Figure S1). In addition, we also detected the higher abundance of the spliced mRNAs in the *rnt1Δ* strain by primer extension analysis using an exon3 specific primer (Figure 3B). Collectively, these results confirm the higher accumulation of spliced *MAT*a1 mRNAs in the *rnt1Δ* strain and show that Rnt1p cleavage competes with splicing, explaining the increased abundance of spliced *MAT*a1 in the absence of Rnt1p.

This technique also allowed us to investigate which exonucleases contribute to the degradation of the unspliced or partially spliced *MAT*a1 transcripts. The largest effect was observed upon inactivation of the nuclear exosome

at 25°C, and of Rat1p in the *rat1-1* strain or in the *xrn1Δ rat1-1* strain at 37°C. The two species that accumulate specifically upon a shift to non-permissive temperature in the *xrn1Δ rat1-1* strain could also be detected by primer extension analysis using an exon3 (Figure 3B) or an intron2 specific primer (Figure 3C), suggesting that they corresponded to fully unspliced and partially spliced species still containing intron2. Sequencing of the RT–PCR products confirmed the identity of these species (data not shown). No accumulation of unspliced or partially spliced species was observed upon inactivation of Xrn1p (Figure 3A) or of Upf1p (Figure S2A), showing that in contrast to many yeast pre-mRNAs (17), unspliced and partially spliced species of *MAT*a1 are not targeted by nonsense-mediated decay (NMD). In addition, inactivation of Xrn1p in the *rat1-1* strain did not result in a major increase of unspliced pre-mRNAs compared to the *rat1-1* stain alone, showing that Rat1p is primarily responsible for the degradation of unspliced and partially spliced *MAT*a1 transcripts. We also found that inactivation of the nuclear exosome in the *rnt1Δ* strain led to an increase of unspliced pre-mRNA, showing that the nuclear exosome can cooperate with Rnt1p in discarding unspliced or incompletely spliced *MAT*a1 species. Overall
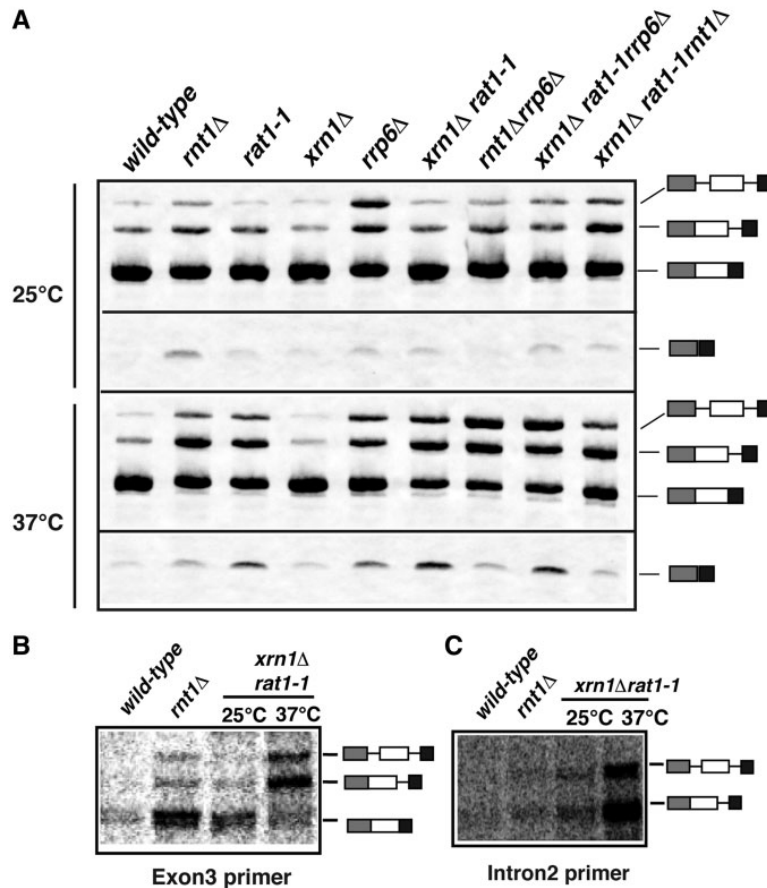
Figure 3. RT PCR and primer extension analysis of *MATa*1 expression in *rnt1Δ* and ribonuclease mutant strains. (A) RT PCR analysis. All panels show RT PCR products obtained using a Cy3-labeled exon1 primer and an unlabeled exon3-primer specific for *MATa*1 in the corresponding strains grown at the indicated temperatures. The exon2-skipped products (shown on the bottom panels) were detected on the same gel as the spliced and unspliced species but that portion of the gel was scanned separately to increase the signal to noise ratio. No other species were detected in the portions of the gel that are not shown. (B) Primer extension analysis of *MATa*1 in the corresponding strains using an oligonucleotide hybridizing to exon3. (C) Primer extension analysis of *MATa*1 in the corresponding strains using an oligonucleotide hybridizing to intron2.

these results show that at least three different nuclear RNA degradation systems can affect unspliced or incompletely spliced *MATa*1 transcripts: Rnt1p, Rat1p and the nuclear exosome.

Exon2-skipped species of *MATa*1 and of other genes containing two introns accumulate in the *rat1-1* and *rrp6Δ* strains

Using the RT–PCR approach described above, we detected a new short molecular weight species migrating faster than fully spliced mRNA, which accumulated specifically in nuclear RNA degradation mutants (Figure 3A, bottom panels). This species accumulated most abundantly at 25°C in the *rnt1Δ* strain (Figure 3A and Supplementary Figure S1A). When shifted to 37°C, this species was found predominantly in the *rat1-1* strain, and in the *rrp6Δ* strain, or in double or triple mutants in which Rat1p is inactivated (Figure 3A). Sequencing of this

product (not shown) revealed that it corresponds to a spliced mRNA in which exon 2 was skipped and where exon 1 was spliced directly to exon 3. This result suggested that these exon2-skipped species are preferentially targeted for degradation by nuclear exonucleases. We did not detect any accumulation of the exon skipped species in a strain carrying a single deletion of *XRN1*, or in the *upf1Δ* strain (Supplementary Figure S2A), showing that it is not subjected to nonsense-mediated mRNA decay, despite its lack of extended open-reading frame.

The accumulation of the exon2-skipped species in the *rnt1Δ* strain at 25°C is unlikely to reflect a direct degradation function for Rnt1p, since the predicted target stem–loop structure is absent in this product. To further test this hypothesis, we overexpressed Rnt1p or a catalytically inactive mutant derivative in the *rat1-1* strain in order to investigate whether Rnt1p overexpression could decrease the levels of the exon2-skipped species that accumulate in the absence of Rat1p activity. The results shown in

Supplementary Figure S3 show that overexpression of Rnt1p does not change the accumulation of exon2-skipped species in the context of the *rat1-1* strain. Thus, it is unlikely that Rnt1p directly cleaves these species. Rather, we favor a model in which Rnt1p cleavage decreases the fraction of precursors that can fold in a conformation that favors the skipping of exon2 (see 'Discussion' section).

Rat1p has been shown to promote transcription termination of RNA polymerase II, both in downstream regions (22) and within the body of genes (23). Association of Rat1p with Rai1p promotes this function, and strains lacking Rai1p are also deficient in transcription termination (22). To investigate if accumulation of the exon2-skipped species in the *rat1-1* strain is an indirect consequence of a transcription termination defect, we assessed exon2-skipped species levels in *rai1Δ* cells. In contrast to the *rat1-1* control, we were unable to detect any accumulation of the exon2-skipped species in samples extracted from the *rai1Δ* strain (Supplementary Figure S2B). We conclude that the accumulation of exon2-skipped species is unlikely to result from indirect transcription termination defects, but more likely to a lack of degradation by Rat1p.

To further characterize why unspliced and exon2 skipped species are preferentially targeted by Rat1p, we investigated whether a lack of polyadenylation would potentially impair their export to the cytoplasm, making them more susceptible to degradation by this nuclear exonuclease. We purified the polyadenylated fraction of RNAs from yeast total RNAs using oligo-dT affinity, and analyzed the enrichment of the different *MATa*1 transcripts in the poly(A)+ fractions (Supplementary Figure S2C). Interestingly, both the various unspliced species and the exon2-skipped species were found in the poly(A)+ fractions of the *rat1-1* strain, suggesting that these species are correctly polyadenylated. This result suggests that a lack of polyadenylation is not the reason why these species are preferentially targeted by Rat1p.

The previous data showed that Rat1p and the nuclear exosome degrade transcripts resulting from mis-splicing events in which the spliceosome had skipped the central exon of *MATa*1. To investigate if this effect is specific to *MATa*1 or can be extended to other genes, we analyzed the expression of other genes containing two introns (*DYN2, SUS1, YOS1*) in the *rat1-1* and *rrp6Δ* strains by RT–PCR. First, we found various unspliced species accumulating in the *rat1-1* and *rrp6Δ* strains (Figure 4), showing that these unspliced RNAs are degraded by Rat1p and the nuclear exosome, as shown above for *MATa*1 and in previous studies for other intron containing genes (24). Figure 4 also shows that an increase of exon2-skipped species of *SUS1, YOS1* and *DYN2* could be detected at permissive temperature (25°C) in the *rrp6Δ* strain. After a shift to 37°C, we also observed a higher accumulation of these exon-skipped species in the *rat1-1* and *rrp6Δ* strains. These results show that exon2-skipped species of several genes accumulate to higher levels when the activity of nuclear exonucleases is disrupted, suggesting that the degradation of these exon skipped species is under the control of nuclear RNA surveillance.
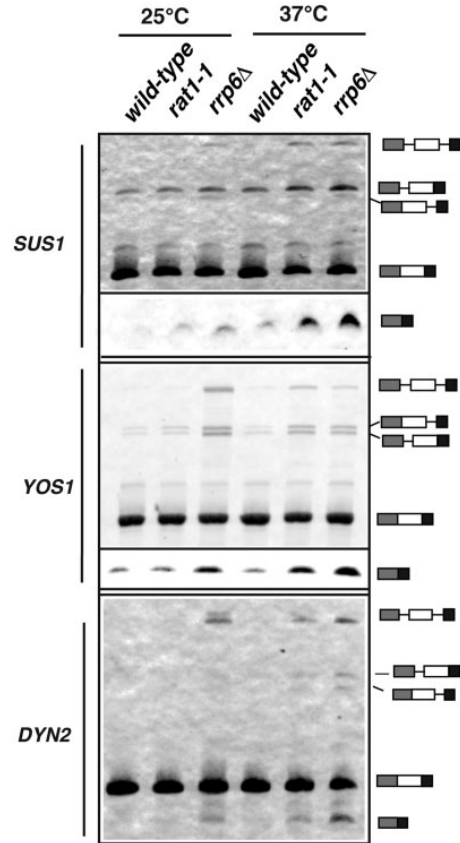

Figure 4. Analysis of exon2 skipping for the *DYN2, SUS1* and *YOS1* genes in the *rat1-1* and *rrp6Δ* strains. All panels show RT PCR products obtained using a Cy3-labeled exon1 primer and an unlabeled exon3-primer specific for the corresponding genes in the strains grown at the indicated temperatures. The exon2-skipped products (shown on the bottom panels) were detected on the same gels as the spliced and unspliced species but that portion of the gel was scanned separately to increase the signal to noise ratio. No other species were detected in the portions of the gel that are not shown.

DISCUSSION

In this study, we used a genetic analysis to investigate the mechanisms of degradation that target the different RNAs produced by the *MATa*1 gene. We found that the unspliced pre-mRNA, as well as a partially spliced version of *MATa*1 that retains intron2 are degraded by Rnt1p, Rat1p and by the nuclear exosome (Figure 3A). As shown previously for the *RPS22B* gene (16), Rnt1p cleavage probably competes with splicing of the pre-mRNA, which explains the increase of abundance of the *MATa*1 mRNA in strains lacking Rnt1p (Figures 1 and 2; Supplementary Figure S1). We note that the terminal loop of the stem–loop structure cleaved by Rnt1p might not be optimal, since it contains the AGNN motif but is a 7-nt loop (Figure 1D) instead of the canonical tetraloop motif (20,25). This suboptimal structure might explain why Rat1p is the major

degradative activity that targets unspliced *MAT*a1 pre-mRNAs, rather than Rnt1p.

We have postulated that the increased abundance of unspliced and partially spliced species in RNA degradation mutants reflect their stabilization. Because these species do not accumulate to significant levels in wild-type cells, we were unable to compare their turnover rates and to estimate their stability in the different backgrounds. Thus, we cannot formally rule out the possibility that the splicing efficiency of *MAT*a1 is influenced by the disruption of degradative activities. For example, it was shown recently that RNA degradation complexes can modulate the rate of other steps in gene expression such as transcriptional elongation (26), and the rate of elongation might affect splicing efficiency. While we cannot formally exclude the possibility of indirect effect for exonuclease mutants, the observation that spliced *MAT*a1 mRNA levels increase in the absence of Rnt1p is consistent with a direct role for Rnt1p in cleaving unspliced mRNAs and a model suggesting that Rnt1p cleavage competes with splicing of the pre-mRNA (Figure 5A).

As opposed to many yeast pre-mRNAs (17), unspliced *MAT*a1 transcripts are not subject to nonsense-mediated mRNA decay (Supplementary Figure S2A), but rely mostly on nuclear degradation pathways. It is possible that these species are not exported very fast, and thus rely mostly on nuclear degradation, as described previously for other genes (24) and for the *DYN2*, *SUS1* and *YOS1* genes (Figure 4). Regardless of the reason why *MAT*a1 seem to rely on nuclear degradation, our results show that the splicing of *MAT*a1 might be suboptimal. This relative inefficient splicing might be caused by the small size of its introns ($\sim$50 nt), which are smaller than most *S. cerevisiae* introns. In addition, folding of the stem–loop structure that sequesters the 5′-splice site of intron2 and prevents U1 snRNP binding (Figure 5A) might also explain the inefficiency of intron2 splicing and why species containing this intron are detected in ribonuclease mutants. Thus, nuclear degradation mechanisms have been selected to limit the accumulation of mRNAs that have escaped the splicing machinery, and also to prevent their expression in the cytoplasm as unfunctional proteins.

Role of the partially spliced *MAT*a1 pre-mRNA encoding a1′

In contrast to most intron containing genes, some partially spliced species of *MAT*a1 encode proteins that do not correspond to truncated forms of the normal protein (Figure 5B). The partially spliced *MAT*a1 mRNA species that accumulates in Rnt1p and Rat1p-deficient strains had been previously analyzed in splicing mutants (9) and further studied by Ner and Smith (8) via experiments on branch point mutants. In this study, it was first tested whether the protein encoded by the *MAT*a1 transcript including Intron2 (termed a1′) could have a biological role and work together with mature a1. This partially spliced transcript would encode a protein that is different and longer than mature a1, especially in the C-terminal domain (Figure 5). The C-terminal domain of
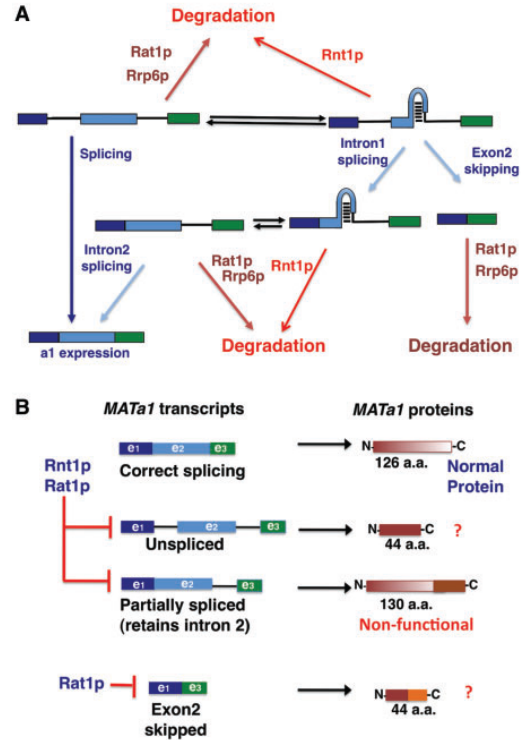

Figure 5. Nuclear RNA quality control discards incompletely or aberrantly spliced forms of *MAT*a1 that result in aberrant a1 isoforms. (**A**) Model of degradation of unspliced, partially spliced and mis-spliced forms of *MAT*a1 by Rnt1p, Rat1p and the nuclear exosome. (**B**) Diagram of the different isoforms of a1 encoded by the different unspliced, partially spliced and mis-spliced forms of *MAT*a1.

a1 is where DNA-binding occurs through Helix 3 when in complex with alpha2. Interactions with alpha2 occur between Helices 1 and 2 and are unaffected in a1′ based on primary sequence and structural analysis (27) (PDB code: 1F43). Specific residues on a1 that are important for contacting DNA are five residues within Helix 3 (4). In their study, Ner and Smith suggested that the a1′ protein does not have any biological role, as it is unable to rescue a1 deficiency (8). However this protein might potentially act as a dominant negative inhibitor, explaining the requirement to limit its accumulation through the degradation of the mRNA encoding this isoform.

Although constitutively expressed in cells of the 'a' mating-type as a part of the haploid *MAT*a transcriptional program, the a1 protein has not yet been characterized to have a direct function in these cells (28,29). In diploid cells, this protein combines with the alpha2 protein in order to repress haploid-specific gene expression. In this study, we have assessed the role of Rat1p and Rnt1p in regulation of the *MAT*a1 transcript in haploid cells. Rnt1p is already known to have important roles in cell-cycle regulation and cell division (30), and is a crucial enzyme for correct processing of numerous non-coding RNAs as well as degradation of mRNAs. Our microarray data

show that one of the genes that are considerably down-regulated (−12.3-fold) in *rnt1Δ* cells encodes HO, the mating-type switching endonuclease gene that is known to be repressed in diploids by the a1 and alpha2 heterodimer (31,32). Considering the accumulation of mature *MATa*1 in *rnt1Δ* strains, as well as partially unspliced forms, one hypothesis is that the excess of a1 could associate with the a2 protein, which has many sequence similarities to alpha2, to act in repressing HO even within haploid cells. Further experiments are required to prove this model.

A novel RNA surveillance mechanism to degrade exon-skipped species

Strikingly, we also found that incorrectly spliced mRNAs are generated from the *MATa*1 locus, in which the exon2 is skipped and exon1 is spliced to exon3. These species are preferentially degraded by Rat1p and the nuclear exosome, but they also accumulate in the absence of Rnt1p. This suggests that formation of the target stem loop structure that bridges exon2 and intron2 favors exon2 skipping for the transcripts that have folded in this structure (Figure 5A). Formation of this structure would sequester the 5′-splice site of intron2 from U1 snRNP binding, promoting a direct splicing of exon1 to exon3 (Figures 1D and 5A). Thus the increase of exon-skipped species in the *rnt1Δ* strain (Figure 3A and Supplementary Figure S1A) might not be due to the fact that Rnt1p directly contributes to degrading exon skipped species, but rather because it eliminates a fraction of the pre-mRNAs which have folded in a manner that promotes the skipping of exon 2 (Figure 5A). In contrast, our data suggest that Rat1p and the nuclear exosome are directly responsible for degrading these exon-skipped species. Strikingly, we found that exon2-skipped species of other genes that contain two introns also accumulate when Rat1p or the nuclear exosome are inactivated (Figure 4). This suggests that nuclear RNA degradation plays a general role in degrading species that have been incorrectly spliced. This type of event is favored by the small size of central exons in yeast, facilitating the recognition of the downstream 3′-splice site in intron2, rather than the correct 3′-splice site of intron1.

We presumed that the increased abundance of the exon2-skipped products when Rat1p or the nuclear exosome are inactivated reflects their stabilization. However, we cannot formally exclude the possibility that inactivation of Rat1p or of the nuclear exosome indirectly affects splicing, and that this change of splicing efficiency is responsible for an increased rate of exon skipping. For instance, the accumulation of this product might reflect functions for Rat1p other than direct degradation, such as promoting transcriptional termination within coding regions (23). Inactivation of Rat1p results in a faster RNA polymerase, and it is known that a faster rate of elongation promotes exon skipping as shown in mammalian cells. However, *RAI1*-deficient cells that are also deficient in termination (22) do not exhibit an accumulation of the exon2-skipped species. This suggests that

accumulation of the exon2-skipped species is likely due to its lack of degradation.

It remains to be understood why these exon-skipped species are preferentially targeted by Rat1p or the nuclear exosome in the nucleus, and not by nonsense-mediated decay. Since Rat1p associates co-transcriptionally with the RNA polymerase (22), it is possible that it acts non-discriminately on non-capped RNAs, and that exon-skipped species are particularly prone to Rat1p-mediated degradation. Inactivation of Xrn1p, the general cytoplasmic exonuclease involved in NMD does not increase the amount of exon2 skipped species, even when Rat1p is inactivated (Figure 3A). This suggests that these exon2 skipped species are not exported to the cytoplasm but are retained in the nucleus. This would explain why these exon-skipped species are also affected by the nuclear exosome. Regardless of the precise mechanism of nuclear retention, the finding that this nuclear RNA degradation pathway targets species that are incorrectly spliced underscores the importance of nuclear RNA surveillance in the quality control of gene expression and identify a novel mechanism of RNA surveillance for incorrect splicing.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1 3.

REFERENCES

1. 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.
2. Johnson,A.D. (1995) Molecular mechanisms of cell-type determination in budding yeast. *Curr. Opin. Genet. Dev.*, **5**, 552–558.
3. Goutte,C. and Johnson,A.D. (1988) a1 protein alters the DNA binding specificity of alpha 2 repressor. *Cell*, **52**, 875–882.
4. Li,T., Stark,M.R., Johnson,A.D. and Wolberger,C. (1995) Crystal structure of the MATa1/MAT alpha 2 homeodomain heterodimer bound to DNA. *Science*, **270**, 262–269.
5. Goutte,C. and Johnson,A.D. (1993) Yeast a1 and alpha 2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. *J. Mol. Biol.*, **233**, 359–371.
6. Phillips,C.L., Stark,M.R., Johnson,A.D. and Dahlquist,F.W. (1994) Heterodimerization of the yeast homeodomain transcriptional regulators alpha 2 and a1 induces an interfacial helix in alpha 2. *Biochemistry*, **33**, 9294–9302.

7. Kohrer,K. and Domdey,H. (1988) Splicing and spliceosome formation of the yeast MATa1 transcript require a minimum distance from the 5′ splice site to the internal branch acceptor site. *Nucleic Acids Res.*, **16**, 9457–9475.
8. Ner,S.S. and Smith,M. (1989) Role of intron splicing in the function of the MATa1 gene of Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **9**, 4613–4620.
9. Miller,A.M. (1984) The yeast MATa1 gene contains two introns. *EMBO J.*, **3**, 1061–1065.
10. Baudin,A., Ozier-Kalogeropoulos,O., Denouel,A., Lacroute,F. and Cullin,C. (1993) A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Res.*, **21**, 3329–3330.
11. Chanfreau,G., Rotondo,G., Legrain,P. and Jacquier,A. (1998) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J.*, **17**, 3726–3737.
12. Egecioglu,D.E., Henras,A.K. and Chanfreau,G.F. (2006) Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. *RNA*, **12**, 26–32.
13. Chu,S., DeRisi,J., Eisen,M., Mulholland,J., Botstein,D., Brown,P.O. and Herskowitz,I. (1998) The transcriptional program of sporulation in budding yeast. *Science*, **282**, 699–705.
14. Petfalski,E., Dandekar,T., Henry,Y. and Tollervey,D. (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.*, **18**, 1181–1189.
15. Longtine,M.S., McKenzie,A. 3rd, Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **14**, 953–961.
16. Danin-Kreiselman,M., Lee,C.Y. and Chanfreau,G. (2003) RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol. Cell.*, **11**, 1279–1289.
17. 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.
18. Lee,C.Y., Lee,A. and Chanfreau,G. (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of S. cerevisiae box C/D snoRNAs. *RNA*, **9**, 1362–1370.
19. Lee,A., Henras,A.K. and Chanfreau,G. (2005) Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. *Mol. Cell.*, **19**, 39–51.
20. Chanfreau,G., Buckle,M. and Jacquier,A. (2000) Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. *Proc. Natl Acad. Sci. USA*, **97**, 3142–3147.
21. Wu,H., Henras,A., Chanfreau,G. and Feigon,J. (2004) Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc. Natl Acad. Sci. USA*, **101**, 8307–8312.
22. Kim,M., Krogan,N.J., Vasiljeva,L., Rando,O.J., Nedea,E., Greenblatt,J.F. and Buratowski,S. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, **432**, 517–522.
23. Jimeno-Gonzalez,S., Haaning,L.L., Malagon,F. and Jensen,T.H. (2010) The yeast 5′-3′ exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Mol. Cell.*, **37**, 580–587.
24. Bousquet-Antonelli,C., Presutti,C. and Tollervey,D. (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell*, **102**, 765–775.
25. Wu,H., Yang,P.K., Butcher,S.E., Kang,S., Chanfreau,G. and Feigon,J. (2001) A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J.*, **20**, 7240–7249.
26. Kruk,J.A., Dutta,A., Fu,J., Gilmour,D.S. and Reese,J.C. (2011) The multifunctional Ccr4-Not complex directly promotes transcription elongation. *Genes Dev.*, **25**, 581–593.
27. Anderson,J.S., Forman,M.D., Modleski,S., Dahlquist,F.W. and Baxter,S.M. (2000) Cooperative ordering in homeodomain-DNA recognition: solution structure and dynamics of the MATa1 homeodomain. *Biochemistry*, **39**, 10045–10054.
28. Haber,J.E. (1998) Mating-type gene switching in Saccharomyces cerevisiae. *Annu. Rev. Genet.*, **32**, 561–599.
29. Zill,O.A. and Rine,J. (2008) Interspecies variation reveals a conserved repressor of alpha-specific genes in Saccharomyces yeasts. *Genes Dev.*, **22**, 1704–1716.
30. Catala,M., Lamontagne,B., Larose,S., Ghazal,G. and Elela,S.A. (2004) Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. *Mol. Biol. Cell.*, **15**, 3015–3030.
31. Sprague,G.F. Jr (1991) Signal transduction in yeast mating: receptors, transcription factors, and the kinase connection. *Trends Genet.*, **7**, 393–398.
32. Roberts,C.J., Nelson,B., Marton,M.J., Stoughton,R., Meyer,M.R., Bennett,H.A., He,Y.D., Dai,H., Walker,W.L., Hughes,T.R. *et al.* (2000) Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science*, **287**, 873–880.