The yeast *RPL9B* gene is regulated by modulation between two modes of transcription termination

Rajani Kanth Gudipati, Helen Neil¹
Frank Feuerbach, Christophe Malabat
and Alain Jacquier*

Institut Pasteur, Unité de Génétique des Interactions Macromoléculaires,
Paris, France

RNA Pol II transcription termination can occur by at least two alternative pathways. Cleavage and polyadenylation by the CPF/CF complex precedes mRNA transcription termination, while the Nrd1 complex is involved in transcription termination of non-coding RNAs such as sno/snrRNAs or cryptic unstable transcripts. Here we show that transcription of *RPL9B*, one of the two genes coding for the ribosomal protein Rpl9p, terminates by either of these two pathways. The balance between these two pathways is modulated in response to the *RPL9* gene copy number, resulting in the autoregulation of *RPL9B* gene expression. This autoregulation mechanism requires a conserved potential stem-loop structure very close to the polyadenylation sites. We propose a model in which Rpl9p, when in excess, binds this conserved 3'-UTR structure, negatively interfering with cleavage and polyadenylation to the benefit of the Nrd1-dependent termination pathway, which, being coupled to degradation by the nuclear exosome, results in downregulation of *RPL9B* gene expression.

The EMBO Journal (2012) 31, 2427–2437. doi:10.1038/emboj.2012.81; Published online 13 April 2012

Subject Categories: chromatin & transcription

Keywords: budding yeast; post-transcriptional regulation; ribosomal protein gene; *Saccharomyces cerevisiae*; transcription termination

Introduction

In eukaryotes, regulation of gene expression can occur at virtually any level, from transcription to processing, nuclear export, RNA stability, translation and protein stability. In metazoans, a large part of the protein diversity is generated at the level of pre-mRNA maturation. It is well recognized that alternative splicing plays a major role in the regulation of gene expression (Keren *et al.*, 2010; Witten and Ule, 2011). Likewise, the well-known involvement of alternative cleavage and polyadenylation sites (poly(A) sites) usage in

gene regulation (see Edwalds-Gilbert (1997) for review) has recently been underscored by the finding that it is a very widespread mechanism important for normal cell development as well as cancer progression (Mayr and Bartel, 2009).

mRNA 3'-end formation involves the RNA polymerase II (Pol II)-associated cleavage and polyadenylation complex (CPF/CF complex), a large complex composed of more than 20 proteins in yeast (see Kuehner *et al.* (2011) for review). The poly(A) site sequences recognized by CPF/CF consist of loose composite consensus sequences (at least four separate degenerated motifs define a poly(A) site in yeast; Graber, 2002). The 3'-end maturation process of mRNAs is initiated by the recognition of cis-acting sequences (the poly(A) sites) on the nascent transcript by components of the CPF/CF complex that triggers endonucleolytic cleavage of the RNA. The 5'-fragment of the RNA is polyadenylated and exported to the cytoplasm, whereas the 3'-fragment is rapidly degraded by the 5'- to 3'-exonuclease Rat1 (Xrn2 in human). In the torpedo model, termination is proposed to occur when Rat1, which attacks the 5'-end generated by the cleavage of the nascent transcript, ultimately catch up with Pol II, resulting in its dissociation from the DNA template (Kim *et al.*, 2004). The efficiency of termination correlates well with the strength of the poly(A) site. 3'-end maturation of the mRNAs by cleavage and polyadenylation is thus intimately coupled to transcription termination. Yet, there is a second mode of transcription termination, which is carried out by another distinct Pol II-associated complex comprised of Nrd1p, Nab3p and Sen1p (Nrd1 complex). Nrd1p and Nab3p are two RNA-binding proteins that recognize GUAA/G and UCUU sequences, respectively, whereas Sen1p is an RNA helicase. The Nrd1 complex is required for the termination and synthesis of stable non-coding RNAs, such as snRNAs and snoRNAs (Steinmetz *et al.*, 2001), as well as unstable pervasive RNAs called CUTs for cryptic unstable transcripts (Arigo *et al.*, 2006; Thiebaut *et al.*, 2006). As for the CPF/CF complex, the Nrd1 complex interacts with the carboxy-terminal domain (CTD) of the large subunit of Pol II (Vasiljeva *et al.*, 2008). Although the actual mechanism of termination by the Nrd1 complex has not been shown, one idea is that the simultaneous binding of the Nrd1 complex to the nascent transcript and to Pol II could trigger the unwinding by Sen1p of the RNA/DNA duplex within the transcriptional R-loop, resulting in the dissociation of Pol II. Moreover, the Nrd1 complex also interacts with the nuclear exosome and couples this mode of transcription termination to 3'- to 5'-exonucleolytic trimming/degradation (Vasiljeva and Buratowski, 2006). In the case of snRNAs and snoRNAs, co-transcriptionally assembled snRNP and snoRNP proteins restrict exonucleolytic trimming, which thus constitutes one of the maturation steps of these RNPs. In the case of CUTs, however, the exonucleolytic degradation of the RNA is not impeded by proteins, resulting in the complete degradation of the transcripts.

*Corresponding author. Institut Pasteur, Unité de Génétique des Interactions Macromoléculaires, 25 rue du Docteur ROUX, CNRS-UMR3525, F-75724 Paris, Cedex 15, France. Tel.: +33 1 40 61 32 05; Fax: +33 1 40 61 34 56; E-mail: alain.jacquier@pasteur.fr

¹Present address: Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), iBiTec-S, Service de Biologie Intégrative et Génétique Moléculaire, F-91191 Gif-sur-Yvette, Cedex, France.

Received: 7 September 2011; accepted: 13 March 2012; published online: 13 April 2012

These RNAs thus exhibit instability, being degraded as soon as their transcription terminates, hence their 'cryptic' nature since they never accumulate in wild-type cells, at least under standard laboratory growth conditions (Wyers *et al*, 2005). Because of the pervasive nature of Nrd1/Nab3 recognition sequences, this degradation-coupled termination mode is thought to constitute a 'default' transcription termination mode able to target undesired or spurious transcripts to degradation.

While the CPF/CF preferentially binds the CTD of the Rpb1p subunit of Pol II in its serine 2 phosphorylated form, Nrd1p binds the same domain when it is phosphorylated on serine 5. This modification, present at the beginning of transcription elongation, decreases, concomitantly with an increase of serine 2 phosphorylation, while Pol II synthesizes the first few hundred nucleotides of RNA (Gudipati *et al*, 2008; Vasiljeva *et al*, 2008). This is consistent with the Nrd1 complex being dedicated to transcription termination of short non-coding RNA. Indeed, very few examples have been reported that suggest the implication of the Nrd1 pathway in the termination of normal mRNAs in wild-type cells. The Nrd1 termination mode has been reported to be able to serve as a 'fail-safe' mRNA transcription termination mechanism in a *rat1* mutant background in which the poly(A) site-associated transcription termination mechanism is deficient or when the main poly(A) site of an mRNA has been deleted (Rondón *et al*, 2009). The Nrd1 mode of termination is then able to efficiently compete with the use of cryptic, weak poly(A) sites. The transcripts resulting from such aberrant events are unstable. However, it has recently been reported that, in a few cases, the Nrd1 termination pathway could be involved in termination of some mRNAs. Indeed, the recent genome-wide mapping of Nrd1-Nab3 crosslink sites has revealed the presence of these factors 3' to several genes (such as *AHP1*, *URA1* and *SNA3*), suggesting a role for the Nrd1 pathway in termination of their transcription (Jamonnak *et al*, 2011). Furthermore, *CTH2* mRNA synthesis has been shown to involve a peculiar pathway. Transcription terminates in a Nrd1-dependent manner, generating a 3'-extended pre-mRNA, subsequently trimmed by the nuclear exosome that pauses at a [GU]_n site before the RNA gets polyadenylated by Pap1, leading to the production of a mature *CTH2* mRNA (Ciais *et al*, 2008).

Here, we show that the ribosomal protein gene *RPL9B* is autoregulated by a novel mechanism that takes advantage of the ability of Pol II to terminate transcription by these two alternative modes, the CPF/CF mode that gives rise to stable mRNAs and the Nrd1 mode that generates unstable transcripts. Indeed, we show that *RPL9B* transcription naturally terminates by these two alternative pathways and is regulated by the modulation in the choice between these two alternate modes. This modulation requires the presence of a conserved sequence, potentially forming a stem-loop structure, which is in very close proximity to the poly(A) sites. These observations suggest a model in which Rpl9p, when present in excess, binds, directly or indirectly, to this conserved hairpin, hampering the recognition of the poly(A) sites to the benefit of the use of downstream Nrd1/Nab3 termination sites, which trigger degradation-coupled transcription termination, hence ensuring gene repression.

Results

***RPL9B* generates, in addition to canonical poly(A) mRNAs, transcripts with the characteristics of CUTs**

Previous genome-wide analyses determined, at the nucleotide level, the 3'-ends of both wild-type poly(A) transcripts (Neil *et al*, 2009; Ozsolak *et al*, 2010) and transcripts from an RNA fraction highly enriched for CUTs (Neil *et al*, 2009). CUTs are normally highly unstable as a result of their Nrd1/Nab3 mode of transcription termination, which is coupled to fast nuclear degradation by the TRAMP-exosome complex (Jacquier (2009) for review). The CUT-enriched fraction was obtained by selecting for nuclear capped RNAs (RNAs co-precipitated with the affinity-purified nuclear cap-binding factor Cbp20p) from cells in which the CUTs are stabilized by hampering TRAMP and nuclear exosome activities (*depleted-Trf4*, *Δrrp6*). For a number of mRNAs, 3'-end tags could be found in the CUT fraction as well as in the poly(A) fraction. The tags recovered in both fractions essentially mapped to the same sites in both samples, suggesting that they simply represented the nuclear and the cytoplasmic fractions of normal mRNAs, respectively. *RPL9B*, however, behaved differently, its 3'-end tags clustering at well distinct positions in the poly(A) and CUT fractions (Figure 1A). The poly(A) tags cluster into two main regions, separated by about 80 nucleotides, thus defining two transcripts, here called #1 and #2 (Figure 1B), corresponding to two main poly(A) sites. The precise mapping of these poly(A) sites, combining the 3'-tags from Neil *et al*, 2009, the direct RNA sequencing tags (DRS tags) from Ozsolak *et al* (2010), as well as 3'-RACE and RNase H mappings performed in this study, are shown in Supplementary Figure S1. In the CUT fraction, however, a different 3'-tag cluster was found yet about 100 nucleotides downstream of the poly(A) sites. When RNAseq (performed according to Nagalakshmi *et al* (2008); Figure 1A) and RNase H mapping (Figure 1B) were used to analyse the 3'-ends of *RPL9B* RNAs in a *Δrrp6* strain, it confirmed that the 3'-end cluster found in the CUT fraction corresponded to an extended form of the *RPL9B* gene transcript (Figure 1B, transcript #3; compare lanes 1 and 2). That this extended form is absent in a wild-type strain, but revealed by compromising the TRAMP-nuclear exosome activities (Neil *et al*, 2009), suggests that it is derived from a Nrd1/Nab3-dependent transcription termination (which is coupled to nuclear exosome degradation). Also consistent with this hypothesis, these extended *RRP6*-sensitive transcripts terminate just after a region of the sequence harbouring a high density of Nrd1 and Nab3 recognition sequences (Figure 1A; green tags) that coincides with a high-density cluster of Nrd1 and Nab3 crosslink sites (Figure 1A), according to Jamonnak *et al*, (2011).

***RPL9p* regulates the levels of endogenous *RPL9B* mRNAs**

Like the majority of the ribosomal proteins in *Saccharomyces cerevisiae*, Rpl9p is encoded by two gene copies, namely *RPL9A* and *RPL9B*. Rpl9p is part of the large (60S) ribosomal subunit, and Rpl9Ap and Rpl9Bp are nearly identical (98.5%). In the absence of the *RPL9A* gene, the *RPL9B* transcripts #1 and #2 accumulate relative to a wild-type strain (Figure 1B; compare lanes 1 and 3). This indicates

Figure 1 Transcription termination of *RPL9B* occurs at multiple positions. **(A)** Graphic representation of the *RPL9B* locus. The blue and red curves, obtained from wild-type and Rrp6p-depleted strain, respectively, indicate the RNA species as determined by deep sequencing of total RNA (Feuerbach *et al*, unpublished). The y-axis scale is the log2 of the number of reads per nucleotide, with the scale maximum set at 1000. The blue and orange vertical bars represent the RNA 3'-ends as determined either by 3'-long SAGE tags (3' tags; blue bars for the poly(A) fraction, orange bars for the CUTs enriched fraction; Neil *et al*, 2009) or by direct RNA sequencing of mRNA 3'-ends (3' DRS; Oszolák *et al*, 2010). The height of the bars is proportional to the number of reads, with the y-axis scale maximum set at 30 for the 3'-tags and 550 for the 3'-DRS data. The red and green bars represent the crosslink data (Jamonnak *et al*, 2011) obtained for Nab3 (GSM791767_Nab_PolyA data set) and Nrd1 (GSM791766_Nrd_Strata_PolyA data set), respectively. The small green rectangles at the bottom denote the potential Nrd1p- and Nab3p-binding site sequences (GUAA/G or UCUU, respectively). The red double arrow indicates the position of the conserved hairpin (see below). The small horizontal bars labelled #1, #2 and #3 indicate the positions of the 3'-ends as determined by the RNase H experiment illustrated in **B**. The vertical arrows point to the RG030 oligonucleotide-directed RNase H digestion used in **B**. The small horizontal arrows indicate the positions of the RG030 and RG055 oligonucleotides, as labelled. The horizontal black bar indicates the position of radiolabelled *RPL9B* 3'-UTR RNA probe. **(B)** Northern blot analysis of endogenous *RPL9B* mRNA in different strains grown in rich medium (lane 1: BY4741, lane 2: LMA1109, lane 3: LMA1122 and lane 4: LMA1265; see Supplementary Table S1). RG030-directed RNase H-digested RNAs were resolved on a denaturing 6% acrylamide gel and then probed with an *RPL9B* 3'-UTR-specific RNA probe that was generated *in vitro* by T7 RNA polymerase transcription. The poly (A) tails were also cleaved off by including oligonucleotide d(T)₁₈ during the RNase H digestion. Different species are marked with numbers #1 to #3 denoted by a vertical bar. *SCR1*, probed with a ³²P labelled oligonucleotide, was used as a loading control. All parts of the figure (samples and molecular weights markers-MW) originate from the same gel. Figure source data can be found with the Supplementary data.

that *RPL9B* expression is autoregulated, the amount of its mRNAs increasing in a $\Delta rpl9A$ strain to compensate for the decrease in the gene copy number. In contrast, transcript #3 behaves just in the opposite way, being greatly diminished in the $\Delta rpl9A$, $\Delta rrp6$ strain relative to a $\Delta rrp6$ strain (compare lanes 2 and 4). Thus, accumulation of transcripts #1 and #2 appears to occur at the expense of transcript #3. Note also that *RPL9A* expression is not affected by the deletion of *RPL9B* (data not shown) and, thus, only one of the two duplicated genes is autoregulated.

The autoregulation of *RPL9B* is mediated by its 3'-UTR

The results described above, which show that transcripts #1 and #2 are regulated in a way opposite to transcript #3, suggest that the autoregulation mechanism might operate by modulating the balance between the CPF/CF-dependent transcription termination, giving rise to stable transcripts, and the Nrd1/Nab3 mode of termination, which generates highly unstable transcripts. According to this model, the *cis*-regulatory elements would be expected to be located in the 3'-end region of the gene in order to modulate transcription termination.

Figure 2 *RPL9B* autoregulation involves its 3'-UTR. **(A)** Graphical representation of the copper reporter (CUP1-3'UTR centromeric plasmid) used to analyse the involvement of the *RPL9B* 3'-UTR for its autoregulation. An exogenous promoter (Pr-TEF) drives the expression of *CUP1*. The 830 nucleotides following the *RPL9B* stopcodon were fused directly downstream the stop codon of *CUP1*, followed by the *CYC1* terminator. The small horizontal arrows mark the position of oligonucleotides, as indicated. **(B)** Ten-fold serially diluted cells of different strains (as indicated on the top; in addition, all strains were deleted of the *CUP1-1* and *CUP1-2* endogenous genes; WT, LMA1166; $\Delta rpl9B$, LMA1141; $\Delta rpl9A$, LMA1419; $\Delta rpl9A/2\mu$ -RPL9A, LMA1419 transformed with a 2- μ plasmid carrying the *RPL9A* gene; $\Delta rrp6$, LMA1952; see Supplementary Table S1) transformed with different plasmids (as indicated on the left; Ø denotes an empty plasmid) were spotted on minimal medium supplemented, or not, with copper at a final concentration of 0.2 mM, as indicated on the right. **(C)** Northern blot analysis of *CUP1*-*RPL9B*-3'UTR fusion mRNA (expressed from the CUP1-3'UTR reporter plasmid) in different strains (labelled as in **B**). RNAs were digested with RNase H using oligonucleotides RG012 and d(T)₁₈, resolved on a 6% denaturing acrylamide gel and probed with the RG013 ³²P-labelled oligonucleotide. The *SCR1* RNA probed with a ³²P-labelled oligonucleotide was used as a loading control. All parts of the figure panel originate from the same gel. Figure source data can be found with the Supplementary data.

To directly test this hypothesis, we used a *CUP1* reporter gene. Cup1p is a metallothionein that confers copper resistance to yeast (Winge *et al*, 1985). When the endogenous *CUP1-1* and *CUP1-2* genes are deleted, the cells are unable to grow on copper-containing media (data not shown). This copper sensitivity can be complemented by transformation with a plasmid harbouring the *CUP1* ORF, expressed from an exogenous TEF promoter and with a *CYC1* transcription terminator (data not shown), as expected for a strain properly expressing Cup1p. However, when the 830 nucleotides downstream region of the *RPL9B* ORF (*RPL9B* 3'-UTR) was inserted between the *CUP1* ORF and the *CYC1* terminator (Figure 2A), the cells were no longer able to grow in the presence of 0.2 mM CuSO₄, showing that an element in this 830 nucleotide sequence was repressing the proper expression of the reporter gene (Figure 2B, CUP-3'UTR plasmid in a WT strain). Deletion analyses showed that the first 308 nucleotides of these *RPL9B* downstream sequences were necessary and sufficient to elicit this repressive effect (data

not shown). We then analysed this *CUP1* construct in different genetic backgrounds. As can be seen on Figure 2B, the repressive effect exerted by the *RPL9B* 3'-sequences on *CUP1* expression is mostly obliterated when the *RPL9* gene copy number is reduced from two to one by either deleting *RPL9A* or *RPL9B*. Note that, in the $\Delta rpl9A$ strain, the endogenous *RPL9B* should still be autoregulated. Yet, the autoregulation of *RPL9B* expression is clearly not able to fully compensate for the *RPL9A* deletion in term of total Rpl9p level since there is still sufficient derepression of the *CUP1*-3'UTR reporter to allow growth on the copper-containing medium, although slightly less efficiently than in the $\Delta rpl9B$ strain. Finally, and as expected, overexpression of Rpl9p from a plasmid fully restores repression of the *CUP1*-3'UTR reporter, as judged by the total inability of the cells to grow on copper-containing medium. Thus, fusing the 3'-UTR of *RPL9B* downstream of a *CUP1* gene is sufficient to recapitulate the autoregulation of *RPL9B* on this reporter gene. In addition, that the *RPL9B* 3'-UTR is sufficient to allow regulation of *CUP1* in response

Figure 3 Transcript #3 terminates by the Nrd1–Nab3 pathway. The figure presents northern blot experiments performed from 5% denaturing acrylamide gels probed with the *RPL9B* 3′-UTR-specific RNA probe as in Figure 1B. (A) The cells were grown in galactose complete medium (YPgal), then shifted to glucose-containing medium (YPD) for 9 h. Total RNAs were treated with RNase H in the presence of the oligonucleotides d(T)₁₈, RG030 and RG055, as indicated. The strains were: WT, W303; $\Delta rrp6$, DLY710; *Pgal-NAB3*, DLY1990; *Pgal-NAB3Δrrp6*, DLY2029 (see Supplementary Table S1; note that the last two strains also carry a *tetO7-NRD1* fusion, but this is not indicated in the figure since, in this experiment, doxycycline was not used and thus the *NRD1* gene is not repressed). (B) The cells were grown at 25°C and, when indicated, shifted to 35°C for 30 min. Total RNAs were treated with RNase H in the presence of the oligonucleotides d(T)₁₈ and RG030 (the addition of oligonucleotide RG055 did not reveal any additional band; data not shown). The strains were: WT, W303, *Mat α*; $\Delta rrp6$, DLY123; *rna14-3*, DLY142; *rna14-3Δrrp6*, DLY171; see Supplementary Table S1. All parts of the figure panel originate from the same gel. Figure source data can be found with the Supplementary data.

to the *RPL9* gene copy number, even though it is expressed from an exogenous TEF promoter, is consistent with the hypothesis that *RPL9B* autoregulation is a post-transcriptional mechanism.

Finally, deletion of *RRP6* partially restores cell growth in the presence of copper (Figure 2B; $\Delta rrp6$ strain), indicating that at least part of the *CUP-3′UTR* transcripts are normally degraded in the nucleus, consistent with the results obtained for the endogenous *RPL9B* gene (Figure 1B, lane 2). Indeed, analysis of the 3′-ends of the *CUP-3′UTR* transcripts by RNase H digestion followed by northern blot (Figure 2C) gave results very similar to those obtained for the endogenous *RPL9B* gene (Figure 1B).

Transcript #3 terminates by the Nrd1–Nab3 pathway

The initial observation that transcript #3 is sensitive to Rrp6p and found only in the CUT-enriched fraction (Neil *et al*, 2009), together with the fact that it terminates shortly after a cluster of Nab3-binding sites and Nrd1/Nab3 crosslink sites (Jamonnak *et al*, 2011), strongly suggested that its transcription is terminated by the Nrd1 pathway rather than the CPF/CF pathway. In order to directly test this hypothesis, we placed *NAB3* under the control of a galactose-inducible promoter (*Pgal-NAB3*) and *NRD1* under the control of a tetracycline-repressible promoter (*P_{TetO7}-NRD1*), either in a wild-type or a $\Delta rrp6$ background. Depletion of Nab3p alone (shift from galactose- to glucose-containing medium for 9 h—Figure 3A) or both Nab3p and Nrd1p (shift from galactose to glucose plus doxycycline-containing medium for 9 h; not shown) both lead to a decrease in the steady state level of transcript #3 (compare lanes 2 and 4 in Figure 3A). This diminution coincided with the appearance of a new band

(labelled #4 on Figure 3A) that was revealed by RNase H digestion of the samples with an additional oligonucleotide (RG055) located downstream of the Nrd1–Nab3-binding sites in the *RPL9B* 3′-UTR. This new band must result from 3′-readthrough transcripts that give rise to a single species upon RNase H cleavage of their 3′-heterogenous distal sequences. Most importantly, band #4 is insensitive to the absence of Rrp6p, fully consistent with it corresponding to long transcripts that terminate in a Nab3p-independent manner.

In order to get an independent confirmation of the involvement of the Nrd1–Nab3 pathway in the transcription termination of transcript #3, we mutated all the putative Nrd1- and Nab3-binding sites (guaa/g, and ucuu sequences according to Carroll *et al*, 2004) within the 3′-UTR of the *CUP-3′UTR* reporter fusion (mutated version named *CUP-ΔBS*; see Supplementary Figure S2 A for the precise nature of the 11 substitutions introduced). Surprisingly, northern blot analysis did not reveal obvious effect of the substitutions on the amount of transcript #3 (compare lanes 2 to 6 or 4 to 8 in Supplementary Figure S2B; the only noticeable effect was the diminution of transcript #1, which can be explained by the fact that one of the substitution in a potential Nab3 site resides precisely at the major poly(A) site of transcript #1 – see Supplementary Figures S1 and S2). Nevertheless, RNase H digestion in the presence of oligonucleotide RG055 revealed a readthrough band (#4) specific for the *CUP-ΔBS* construct (lanes 14 and 16 in Supplementary Figure S2B). This shows that a fraction of the transcripts escape Nrd1–Nab3 termination at site #3 in the mutant. These readthrough transcripts, that give rise to band #4, are reduced upon deletion of *RPL9A*, along with transcript #3, consistent with the idea that they constitute readthrough of the latter. Of note,

band #4 is stabilized by the deletion of *RRP6* (in contrast to band #4 in Figure 3A), which means that the corresponding readthrough transcripts likely terminate at downstream Nrd1–Nab3 sites rather than at a cryptic poly(A) site. The observation that the effect of the substitutions is so weak likely illustrates the fact that the actual Nrd1- and Nab3-binding site consensus sequences remain very poorly defined and is consistent with similar observations made on CUTs (FF and AJ, unpublished results).

Finally, we analysed how mutations in components of the CPF/CF complex affected the different *RPL9B* transcripts species. Mutations in *RNA14* or *RNA15* have been shown to lead to a transcriptional readthrough of termination sites (Mandart and Parker, 1995), the resulting extended RNAs being usually rapidly degraded in the nucleus (Libri *et al*, 2002). We thus analysed the effects of the temperature sensitive (*ts*) mutants *rna14-3* and *rna15-2* (Minvielle-Sebastia *et al*, 1991) on the *RPL9B* transcript profile, either in an *RRP6* wild-type or an $\Delta rrp6$ background. Both *rna14-3* and *rna15-2* exhibited identical results (Figure 3B and data not shown). At the permissive temperature (25°C), the *rna14-3* mutation already exerted a marked effect on cleavage and polyadenylation, with significant readthrough of poly(A) site #1, resulting in a significant decrease in the amount of the corresponding transcript and a concomitant increase of transcript #2 (compare lanes 1 and 2). This suggests that this mutation, at 25°C, impairs termination at site #1 without significantly affecting termination at site #2. After a 30 min shift to the semi-restrictive temperature (35°C), both transcripts #1 and #2 are decreased (compare lanes 5 and 6) and part of the termination probably occurs further downstream in a Nrd1–Nab3-dependent fashion, as evidenced by the increase of transcript #3 in the $\Delta rrp6$ background (compare lanes 7 and 8) and by the fact that no band corresponding to longer transcripts (such as band #4 in Figure 3A) was revealed after RNase H digestion in the presence of oligonucleotide RG055 (not shown).

Altogether, these data confirm the hypothesis that transcript #3 is terminated by the Nrd1–Nab3 pathway, while transcripts #1 and #2 are terminated by the normal pathway for mRNAs, that is, the CPF/CF pathway.

A potential conserved stem-loop structure within the 3'-UTR is essential for the autoregulation

In order to identify functional elements within the *RPL9B* 3'-UTR, we searched for sequences that would be conserved among related yeast species. This revealed a region of conservation, located between about 155 to 200 nucleotides from the end of the ORF (Figure 4A), which has the potential to form a stable hairpin (Figure 4B). Interestingly, many nucleotide differences between yeast species (in red in Figure 4B) exhibit compensatory base changes, a signature for the functional nature of this potential structure. Moreover, and importantly, deletion of this conserved sequence (see Figure 4C for the precise nature of the mutation) in the *CUP1* reporter construct (*CUP1*- Δ hairpin) abrogates the repressive effect of the *RPL9B* 3'-UTR, allowing proper Cup1p expression and growth on the copper-containing medium (data not shown) and totally abolishes the ability of the *CUP1* reporter transcript to respond to *RPL9* gene copy number variations (Figure 4D and data not shown). RNase H mapping of 3'-ends of the corresponding mRNAs shows that deletion of this

potential hairpin generates a major stable transcript whose 3'-end coincides with that of the endogenous *RPL9B* transcripts #1, as determined by 3'-RACE (Supplementary Figure S1). This shows that deletion of the potential hairpin not only makes the gene refractory to Rpl9p-dependent repression, but also transforms the weak poly(A) site of transcript #1 into a strong poly(A) site.

To further analyse the sequence elements involved in the feedback repression by the 3'-UTR, we carried out random PCR mutagenesis of the *RPL9B* 3'-sequences. The mutagenized *RPL9B* 3'-UTRs were cloned between the *CUP1* ORF and the *CYC1* terminator in the reporter construct (see Figure 2A) and selected for their ability to restore growth of *cup1-1 cup1-2* cells on copper-containing media with different molar concentrations of CuSO_4 . The selected clones were sequenced to determine the specific mutations that abrogated the repression and only the clones bearing a single mutation were taken into account. With the exception of a U to G mutation, four nucleotides upstream of the hairpin, 11 out of the 12 mutations that most strongly abrogate the repression (in red in Figure 4C) are located within the conserved stem-loop structure. Mutations exhibiting weaker phenotypes were also found within the hairpin and up to 29 nucleotides upstream (in orange for the mutants moderately affected and in light blue for the mutants weakly affected, as determined by growth in the presence of various CuSO_4 concentrations; data not shown). For some representative mutants, we analysed the steady state levels of the reporter transcripts by RNase H-mediated 3'-end analysis to determine whether the loss of repression correlates with changes in the abundance of reporter transcripts (Figure 4E). This gave rise to a more complex picture than expected. In some of the mutants, a simple increase in the amount of transcripts #1 and #2 was observed, as previously observed in the $\Delta rpl9A$ strain (Figure 1B). This is in particular true for nucleotides that are not involved in Watson–Crick base pairing, such as mutants number 2 or 12. This is the behaviour expected for mutants that simply affect binding of the hairpin by a trans-acting factor (possibly Rpl9p; see below). In contrast, the other mutants reveal an additional termination site, located a few nucleotides upstream of the 3'-end of transcript #1 (labelled with a star on Figure 4E). While this novel 3'-end was not mapped at the nucleotide level, the RNase H experiment locates it within a few nucleotides upstream of the 3'-end of transcript #1, indicating that this new transcript terminates in the hairpin. Strikingly, all these mutants, with the exception of mutant number 11, disrupt a Watson–Crick base pair within the structure. Note that mutant 9, which, while involving a paired nucleotide, is conservative with respect to base pairing, substituting a G:C base pair by a G:U base pair, does not induce this new 3'-end. This suggests that the destabilization of the hairpin might unmask a cryptic poly(A) site, not normally accessible. This effect on the derepression of the *CUP1* reporter might thus be somewhat indirect. This is true for example for mutants 6, 7 and 10, which, apart from this new 3'-end, do not show a marked increase of transcripts #1 and #2. Yet, many mutants, such as mutants 3 and 4, show both the new 3'-end and also a strong increase of transcripts #1 and #2, suggesting that the corresponding nucleotides have a dual role, in maintaining the integrity of the structure and also allowing the recognition by a trans-acting factor (although we are aware that the two

effect are probably not independent from each other). In these examples, the mutation might not totally destabilize the hairpin, which thus would be in equilibrium between a folded and an unfolded state. In the folded state, the mutation might impair the binding of the trans-acting factors, while in the unfolded state, it would unmask the cryptic poly(A) site. Only mutant number 11 does not fit well in this picture.

It does very significantly increase the amount of transcripts #1 and #2, suggesting that it is involved in the recognition by a trans-acting factor, yet it is not well conserved. Moreover, it induces the formation of the new 3'-end, suggesting that it affects the stability of the hairpin, yet it is located outside the stem-loop structure. Several hypotheses could explain these observations. For example, this mutation could favour

alternative structures, competing with the formation of the hairpin. Folding programmes could not reveal obvious competing structures, but these predictions are not fully reliable, not taking into account tertiary interactions for example. Finally, weak mutations, located outside of the hairpin (number 5 and 8), also increase transcripts #1 and #2 (although more modestly than, for example, mutant 2, as expected from their weaker phenotype). This could suggest that the *cis*-regulatory elements recognized by trans-acting factor(s) might extend significantly upstream of the hairpin, although this region is outside the conserved sequences. Here again, other explanation could explain the phenotype. For example, these mutations could affect the overall stability of transcripts #1 and #2 in the cytoplasm, in a manner independent of the amount of Rpl9p. Clearly, more detailed work will be required to fully decipher all the parameters involved in the regulation. Yet, these experiments show that very few sequences, if any, are involved in the repression outside the hairpin, the deletion of which totally abrogates the regulation and where the vast majority of the single mutations were recovered.

Discussion

In this work, we show that *RPL9B* transcription termination occurs by two alternative pathways, the CPF/CF pathway, which gives rise to normal polyadenylated mRNAs, and the Nrd1 pathway, which is coupled to rapid nuclear degradation involving the nuclear exosome factor Rrp6p. We further show that *RPL9B* expression is autoregulated and that this feedback loop involves a shift in the balance between the uses of the two alternative modes of termination. We show that fusing the 3'-UTR region of *RPL9B* to a *CUP1* reporter transcript is sufficient to recapitulate this regulation on *CUP1* expressed from an exogenous promoter, which is fully consistent with this autoregulation being a post-transcriptional process. We finally show that the *RPL9B* 3'-UTR harbours an evolutionary conserved sequence potentially forming a stem-loop structure, which is essential for the regulation and which overlaps or is very close to the two poly(A) sites.

From these observations, we would like to propose the model depicted in Figure 5. The key feature of the model is that the autoregulation of *RPL9B* expression results from the modulation of the balance between a productive mode of transcription termination, which involves the normal CPF/CF mRNA termination pathway, and the degradation-coupled Nrd1/Nab3-dependent mode of termination, normally associated with the transcription of non-coding RNAs such as CUTs or sn/snoRNAs. Two elements are key in the model, a *cis*-acting element, the conserved hairpin, which overlaps or is very close to the poly(A) site recognition sequences, and a trans-acting factor that, by interacting with the hairpin, inhibits the recognition of the poly(A) site signal sequences by the CPF/CF complex. At this stage, the most parsimonious hypothesis is that this trans-acting factor is Rpl9p itself, which has the required property of being an RNA-binding protein. This model is reminiscent of the mode of regulation previously proposed for *RPS28B* (Badis *et al*, 2004). Rps28p binds a conserved hairpin within the 3'-UTR of its mRNA, which promotes the cytoplasmic mRNA decapping and degradation of the RNA in an Edc3-dependent manner. This model is also reminiscent of the mode of regulation of *NPL3*

(Lund *et al*, 2008). The *NPL3* locus produces at least three different RNA species with varying 3'-UTR lengths. The presence of an excess of phosphorylated Npl3p promotes the use of the most downstream poly (A) site, which results in the degradation of the transcripts by both a cytoplasmic and a nuclear degradation pathway, although not involving the Nrd1/Nab3/Sen1 complex.

Despite repeated efforts by various strategies, we were unable to biochemically demonstrate the interaction between the conserved hairpin and Rpl9p. Three-hybrid did not give convincing results either. This suggests that the interaction between Rpl9p and the *RPL9B* 3'-UTR is transient or that factors other than Rpl9p are required for this interaction to occur, similar to what has been proposed for Npl3p (Lund *et al*, 2008). Note also that, as for Rps28p, we could not find, within the region of the 25S rRNA secondary structures bound by Rpl9p (as determined from the crystal structure; Ben-Shem *et al*, 2010), any convincing resemblance with the conserved RNA hairpin within the *RPL9B* 3'-UTR. The same observation was made for Rps14B, which has been shown to directly bind a conserved structure of its own mRNA, structure that does not bear any obvious similarities with its known ribosomal RNA-binding RNA structure (Fewell and Woolford, 1999).

Whatever the trans-acting binding factor is, another parameter that might be essential for such a mechanism to be functional is its ability to efficiently compete, kinetically speaking, with the recognition of the poly(A) site elements by the CPF/CF machinery. The conserved hairpin might simply be located sufficiently upstream of the poly(A) site, such that, being transcribed significantly before the poly(A) site recognition sequences, it would be bound by the trans-acting factor before cleavage and polyadenylation happen. But then, the bound protein might be too far to interfere, by simple steric hindrance for example, with poly(A) site recognition. Thus, the model must accommodate two seemingly contradictory parameters, first the close proximity between the conserved structure and the poly(A) site, second the competition between this factor and CPF/CF for binding to their respective target sequences. Indeed, the hairpin overlaps poly(A) site #1 recognition sequences (a potential Positioning Element, or Element 2, of poly(A) site #1, AAUAAA, resides within the loop of the hairpin), while the most probable Element 1 of poly(A) site #2 (UGCAUA according to the probabilistic model; Graber, 2002) is separated from the hairpin by only 35 nucleotides, which could be sufficient to allow interference between Rpl9p and CPF/CF binding. One could thus assume that the competing poly(A) site should be intrinsically weak. Since we have shown that transcription of *RPL9B* terminates at two poly(A) sites, the first one, which is the closest to the potential hairpin, must be weak to allow significant read through down to the second site. According to a probabilistic poly(A) site prediction model (Graber, 2002), site #1 would indeed be a very weak site, while site #2 would be stronger. This is also confirmed by the effect of the *ma14-3* mutation, which is stronger on site #1 than on site #2 (see Figure 3B). Independent experimental data from the literature suggest that both *RPL9B* poly(A) sites are intrinsically weak. Indeed, in very large-scale Synthetic Genetic Array screens (Costanzo *et al*, 2010), Δ rpl9A was found among the mutations that exhibit the strongest synthetic genetic interaction with two

Figure 5 Proposed model for the autoregulation of *RPL9B*. The RNA is indicated in grey, whereas the potential RNA stem-loop region is indicated by a black double-headed line curve on the DNA, depicted as a black horizontal line. The bold, red question mark points to the fact that we were unable to show the direct interaction between Rpl9p and the RNA stem-loop structure, suggesting that this interaction might be indirect.

thermosensitive mutants of Pcf11p, a key factor of the CPF/CF complex (Supplementary Figure S3). This observation can be readily explained by the fact that the *RPL9B* poly(A) sites are particularly sensitive to the *pcf11* mutations, even in the permissive conditions used during the screens. In the absence of *RPL9A*, the cells can grow well with the *RPL9B* gene alone, given that it is not repressed. In the presence of the *pcf11* mutations, however, *RPL9B* is not properly expressed, most likely as a result of a 3'-end termination defect, and can no longer compensate for the absence of *RPL9A*. In contrast, *RPL9A* cleavage and polyadenylation does not seem to be sensitive to the *pcf11* mutations in the same condition since the ability of *RPL9A* to compensate for the absence of *RPL9B* is not affected by these mutants, which even exhibit weak epistasis with the $\Delta rpl9B$ deletion (Supplementary Figure S3).

The proposed mode of regulation of *RPL9B* is also reminiscent of the autoregulation of yet another RNA-binding protein, Nab2p. In this system, the overexpression of Nab2p results in the destabilization of its own mRNA in a way dependent on an unusual encoded terminal A₂₆

sequence and Rrp6p. Although the precise mechanism by which this regulation occurs remains to be elucidated, it has also been proposed that it might act at the level of mRNA 3'-end formation at an intrinsically weak poly(A) site (Roth *et al*, 2009).

RPL9B represents the fifth example of an autoregulated ribosomal protein gene in *S. cerevisiae*. It is noteworthy that the four other known examples also involve post-transcriptional mechanisms operating at the RNA level: *RPS14B* (Li *et al*, 1995) and *RPL30* (formerly L32) are autoregulated at the level of pre-mRNA splicing (Dabeva and Warner, 1993). The autoregulation of *RPL30* also occurs at the translation level. *RPS3* (Hendrick *et al*, 2001), *RPL4A* (formerly ribosomal protein L2; Presutti *et al*, 1995) and *RPS28B* (Badis *et al*, 2004) are autoregulated at the level of mRNA decay. In this context, the autoregulation mechanism of *RPL9B*, which involves the modulation in the balance between two modes of transcription termination, appears quite original.

Finally, it has previously been shown that the Nrd1/Nab3 mode of termination is most efficient at the beginning

of transcription, when the C-terminal heptameric repeats (CTD) of the RNA Pol II largest subunit are still phosphorylated at the Ser5 positions. It has experimentally been demonstrated that the progressive substitution of Ser5 phosphorylation by Ser2 phosphorylation of the CTD results in the Nrd1-dependent pathway becoming very inefficient past the first 900 nucleotides of transcription (Gudipati *et al*, 2008; Vasiljeva *et al*, 2008). This suggests that the specific mode of regulation described here, relying on the modulation of the balance between the two modes of transcription termination, can only apply to relatively short transcripts, in a range where both the CPF/CF and the Nrd1 pathway can efficiently operate. Indeed, with a size of slightly more than 800 nucleotides, the *RPL9B* transcription unit might have the ideal size to be controlled by such a mechanism.

This raises the question of whether other genes are regulated by the same mechanism. Searching the data from Neil *et al* (2009), we could not identify obvious candidates with clearly separated CPF/CF- and Nrd1-dependent termination sites. But *RPL9B* might have been in the ideal situation where it was expressed at an intermediate level of regulation, thus generating tags in both RNA populations.

Materials and methods

Construction and manipulation of yeast strains and plasmids

All the strains used in this study are derived from W303 or BY4741 and are listed in Supplementary Table S1. The strains with names beginning with DLY are gifts from D Libri. Strains with names starting with LMA are from this work or previous work from our laboratory. To obtain the plasmid borne *CUP-RPL9B-3'UTR* constructs, we first cloned the *CUP1-Tcycl* gene under the control of the promoter TEF (translation elongation factor) generating plasmid *P_{TEF}-CUP1-T_{CYC1}*. *RPL9B 3'-UTR* (832 nt) was subsequently introduced by homologous recombination using NotI linearized plasmid *P_{TEF}-CUP1-T_{CYC1}* and PCR fragment generated by oligos RG005/RG006. All constructs were verified by sequencing.

Random mutations in the *RPL9B 3'-UTR* were generated by random mutagenesis PCR, as described (Wilson and Keefe, 2001), using oligos RG05/RG006. The amplified fragments were cloned by homologous recombination as described above. The mutations were selected by picking the clones that were able to grow in the presence of copper and subsequently sequenced. Only the clones that have a single mutation are represented.

References

- Arigo JT, Eyler DE, Carroll KL, Corden JL (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell* **23**: 841–851
- Badis G, Saveanu C, Fromont-Racine M, Jacquier A (2004) Targeted mRNA degradation by deadenylation-independent decapping. *Mol Cell* **15**: 5–15
- Ben-Shem A, Jenner L, Yusupova G, Yusupov M (2010) Crystal structure of the eukaryotic ribosome. *Science* **330**: 1203–1209
- Carroll KL, Pradhan DA, Granek JA, Clarke ND, Corden JL (2004) Identification of cis elements directing termination of yeast non-polyadenylated snoRNA transcripts. *Mol Cell Biol* **24**: 6241–6252
- Ciais D, Bohnsack MT, Tollervey D (2008) The mRNA encoding the yeast ARE-binding protein Cth2 is generated by a novel 3' processing pathway. *Nucleic Acids Res* **36**: 3075–3084
- Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JLY, Toufighi K, Mostafavi S, Prinz St J, Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M *et al* (2010) The genetic landscape of a cell. *Science* **327**: 425–431
- Dabeva MD, Warner JR (1993) Ribosomal protein L32 of *Saccharomyces cerevisiae* regulates both splicing and translation of its own transcript. *J Biol Chem* **268**: 19669–19674

Yeast culture parameters

The thermo-sensitive mutants were heat shocked, when required, by adding one volume of culture at 25°C into one volume of fresh medium at 35°C and then incubated at 35°C for 30 min. The strains containing plasmids were always grown in minimal medium.

To perform the metabolic depletion of Nab3p, strains carrying the *P_{GAL}-NAB3* allele were grown in galactose-containing rich medium and then shifted for 9 h into glucose-containing rich medium.

RNA preparations and analyses

RNAs were analysed on 5 or 6% denaturing polyacrylamide. Hybridizations were performed using a commercial buffer (Ultrasch, Ambion or Rapid-hyb Buffer, GE healthcare). Strand-specific probes were prepared by standard methods, either by phosphorylation of oligonucleotides using ³²P γATP or by *in vitro* transcription of RNA by including ³²P αUTP using T7 RNA polymerase. The templates for T7 RNA polymerase were prepared by standard PCR, by integrating T7 promoter into one of the PCR primers. RNase H digestions were performed as described (Libri *et al*, 2002), with oligonucleotides as indicated in the figure legends. See Supplementary Table S2 for the sequence of the oligonucleotides used in this study.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank the members of the GIM laboratory for helpful discussions, D Libri for the kind gift of strains (DLY strains in Supplementary Table S1) and Laurence Ma from the Institut Pasteur Genomics Platform (PF1) for running the cDNA deep-sequencing samples. This work was supported by grant ANR 'CutReg' (ANR-08-BLAN-0038) from the Agence Nationale de la Recherche, by the CNRS and the Institut Pasteur. RKG received a 'Roux' fellowship from Institut Pasteur. HN was supported by grant ARN 'CUT' (ANR-ANR-05-BLAN-0062).

Author contribution: RKG, HN and FF constructed the plasmids and yeast strains. RKG and FF performed the RNA analyses and genetic screens, FF prepared the samples for cDNA deep sequencing and CM performed the bioinformatics analyses. RKG, HN, FF and AJ designed the experiments. RKG and AJ wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

- Edwards-Gilbert G (1997) Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* **25**: 2547–2561
- Fewell SW, Woolford Jr JL (1999) Ribosomal protein S14 of *Saccharomyces cerevisiae* regulates its expression by binding to RPS14B pre-mRNA and to 18S rRNA. *Mol Cell Biol* **19**: 826–834
- Graber JH (2002) Probabilistic prediction of *Saccharomyces cerevisiae* mRNA 3'-processing sites. *Nucleic Acids Res* **30**: 1851–1858
- Gudipati RK, Villa T, Boulay J, Libri D (2008) Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* **15**: 786–794
- Hendrick JL, Wilson PG, Edelman II, Sandbaken MG, Ursic D, Culbertson MR (2001) Yeast frameshift suppressor mutations in the genes coding for transcription factor Mbf1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis. *Genetics* **157**: 1141–1158
- Jacquier A (2009) The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet* **10**: 833–844
- Jamonnak N, Creamer TJ, Darby MM, Schaughency P, Wheelan SJ, Corden JL (2011) Yeast Nrd1, Nab3, and Sen1 transcriptome-wide

- binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA* **17**: 2011–2025
- Keren H, Lev-Maor G, Ast G (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* **11**: 345–355
- Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedeia E, Greenblatt JF, Buratowski S (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**: 517–522
- Kuehner JN, Pearson EL, Moore C (2011) Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol* **12**: 283–294
- Li Z, Paulovich AG, Woolford Jr JL (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**: 6454–6464
- Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH (2002) Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol* **22**: 8254–8266
- Lund MK, Kress TL, Guthrie C (2008) Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. *Mol Cell Biol* **28**: 3873–3881
- Mandart E, Parker R (1995) Effects of mutations in the *Saccharomyces cerevisiae* RNA14, RNA15, and PAP1 genes on polyadenylation *in vivo*. *Mol Cell Biol* **15**: 6979–6986
- Mayr C, Bartel DP (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**: 673–684
- Minvielle-Sebastia L, Winsor B, Bonneaud N, Lacroute F (1991) Mutations in the yeast RNA14 and RNA15 genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. *Mol Cell Biol* **11**: 3075–3087
- Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**: 1344–1349
- Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**: 1038–1042
- Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* **143**: 1018–1029
- Presutti C, Villa T, Hall D, Pertica C, Bozzoni I (1995) Identification of the cis-elements mediating the autogenous control of ribosomal protein L2 mRNA stability in yeast. *EMBO J* **14**: 4022–4030
- Rondón AG, Mischo HE, Kawauchi J, Proudfoot NJ (2009) Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Molecular Cell* **36**: 88–98
- Roth KM, Byam J, Fang F, Butler JS (2009) Regulation of NAB2 mRNA 3'-end formation requires the core exosome and the Trf4p component of the TRAMP complex. *RNA* **15**: 1045–1058
- Steinmetz EJ, Conrad NK, Brow DA, Corden JL (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* **413**: 327–331
- Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. *Mol Cell* **23**: 853–864
- Vasiljeva L, Buratowski S (2006) Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol Cell* **21**: 239–248
- Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15**: 795–804
- Wilson DS, Keefe AD (2001) Random mutagenesis by PCR. *Curr Protoc Mol Biol* Chapter 8: Unit 8.3: 1–9
- Winge DR, Nielson KB, Gray WR, Hamer DH (1985) Yeast metallothionein. Sequence and metal-binding properties. *J Biol Chem* **260**: 14464–14470
- Witten JT, Ule J (2011) Understanding splicing regulation through RNA splicing maps. *Trends Genet* **27**: 89–97
- Wyers F, Rougemaille M, Badis G, Rousselle J-C, Dufour M-E, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, Libri D, Jacquier A (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**: 725–737