Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast

Marc Larochelle¹, Marc-Antoine Robert², Jean-Nicolas Hébert¹, Xiaochuan Liu³, Dominick Matteau², Sébastien Rodrigue², Bin Tian³, Pierre-Étienne Jacques²,4 & François Bachand¹,4

Termination of RNA polymerase II (RNAPII) transcription is a fundamental step of gene expression that is critical for determining the borders between genes. In budding yeast, termination at protein-coding genes is initiated by the cleavage/polyadenylation machinery, whereas termination of most noncoding RNA (ncRNA) genes occurs via the Nrd1–Nab3–Sen1 (NNS) pathway. Here, we find that NNS-like transcription termination is not conserved in fission yeast. Rather, genome-wide analyses show global recruitment of mRNA 3′ end processing factors at the end of ncRNA genes, including snoRNAs and snRNAs, and that this recruitment coincides with high levels of Ser2 and Tyr1 phosphorylation on the RNAPII C-terminal domain. We also find that termination of mRNA and ncRNA transcription requires the conserved Ysh1/CPSF-73 and Dhp1/XRN2 nucleases, supporting widespread cleavage-dependent transcription termination in fission yeast. Our findings thus reveal that a common mode of transcription termination can produce functionally and structurally distinct types of polyadenylated and non-polyadenylated RNAs.
RNA polymerase II (RNAPII) is responsible for the synthesis of a broad range of coding and noncoding transcripts, and termination pathways have a decisive influence on the fate of transcribed RNAs. Although transcription termination has been studied in a wide range of organisms, it has been most extensively studied in the model organism *Saccharomyces cerevisiae*, where two major termination pathways exist depending on the class of genes transcribed. For protein-coding genes, data support a mechanism triggered by the cleavage activity of the Ysh1 endonuclease (CPSF-73 in humans), which is part of a nucleosome module within the larger mRNA 3’ end processing complex. Specifically, the co-transcriptional recruitment of conserved cleavage and polyadenylation factors at poly(A) signal (PAS) causes cleavage of the nascent pre-mRNA, followed by 3’ end polyadenylation of the released transcript by the poly(A) polymerase. This cleavage also provides a free and uncapped 5’ entry point for a protein complex that includes an evolutionarily conserved 5’–3’ exonuclease (XRN2 in humans, Rat1 in *S. cerevisiae*; Dhp1 in *S. pombe*). The exonuclease is thought to chase RNAPII and promote its dissociation from the DNA template, a mechanism referred to as “torpedo”-mediated transcription termination.

In addition to mRNAs, RNAPII synthesizes an extensive set of noncoding RNAs (ncRNAs) that include small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), long noncoding RNAs (lncRNAs), and cryptic unstable transcripts (CUTs). In *S. cerevisiae*, transcription of snoRNAs, snRNAs, and CUTs does not rely on the cleavage/polyadenylation machinery for termination, but instead uses a cleavage-independent termination pathway that requires a complex (referred to as NNS complex) consisting of the RNA-binding proteins Nab3 and Nrd1 as well as the DNA/RNA helicase Sen1. In this mode of termination, the NNS complex interacts with both the transcription machinery and specific RNA motifs enriched downstream of ncRNA genes to engage the transcription elongation complex via the Sen1 helicase, which translocates onto the nascent RNA and catches up with the transcribing polymerase to elicit termination.

One transcription-associated feature that appears to influence the choice between torpedo- and NNS-mediated termination is the phosphorylation status of the carboxy-terminal domain (CTD) of the RNAPII catalytic subunit, Rpb1. The CTD consists of a succession of conserved heptad repeats, with the consensus amino acid sequence Y-S-P-T-S-P-S. The RNAPII CTD is subjected to a plethora of post-translational modifications throughout the transcription cycle that are key to coordinate the sequential recruitment of RNA-processing factors. Among CTD modifications, phosphorylation of Ser2 and Ser5 have been most extensively studied and appear to be the most abundant in *S. cerevisiae*. In the cases of small ncRNA genes, NNS recruitment is influenced by Ser5 phosphorylation (Ser5-P) via the CTD interaction domain (CID) of Nrd1. As transcription elongation progresses into protein-coding genes, the ratio between the levels of phosphorylated Ser2 (Ser2-P) and Ser5-P gradually increases, peaking near the mRNA cleavage site and CTD domain (CID) of Pcf11, a component of the mRNA 3’ end processing machinery, preferentially recognizes Ser2-P CTD repeats via its CID domain, which may explain the robust levels of Pcf11 at the 3’ end of protein-coding genes. Yet, Pcf11 recruitment at the 3’ end of mRNA genes may also be influenced by additional CTD modifications. For instance, Tyr1 phosphorylation (Tyr1-P) can impair binding of Pcf11 to Ser2-P CTD peptides in vitro. Accordingly, it was proposed that Tyr1-P prevents recruitment of termination factors in the coding region of genes in *S. cerevisiae*, and that timely dephosphorylation of Tyr1-P upstream of the polyadenylation site by the Gicl7 phosphatase would allow for recruitment of the 3’ end processing machinery via recognition of Ser2-P CTD repeats by the CID domain of Pcf11.

High-throughput sequencing analyses indicate that the majority of genomes are transcriptionally active, thus yielding a large amount of ncRNAs. Yet, how RNAPII transcription is terminated at ncRNA genes remains poorly understood for many eukaryotic species. Also unclear is whether the use of distinct pathways to terminate coding and noncoding transcription, such as described in *S. cerevisiae*, is evolutionarily conserved. Recently, we have shown that the *S. pombe* homolog of *S. cerevisiae* Nrd1, Seb1, does not physically associate with Nab3 and Sen1 homologs in fission yeast. Here we show that the absence of Nab3 and Sen1 homologs in *S. pombe* does not affect transcription termination of coding and ncRNA genes. Unexpectedly, we find that mRNA 3’ end processing factors are enriched at the 3’ end of ncRNA genes and that this recruitment coincides with high levels of Ser2 and Tyr1 CTD phosphorylation. Consistently, we find that most independently transcribed fission yeast snoRNA genes are cleaved and polyadenylated in a manner dependent on conserved 3’ end processing factors, and that termination at ncRNA genes is sensitive to deficiencies in Ysh1 and Dhp1. Our findings indicate that torpedo-mediated transcription termination is widespread in fission yeast, thereby revealing that a universal mode of transcription termination can promote the synthesis and accumulation of both polyadenylated mRNAs and non-polyadenylated ncRNAs.

### Results

**Nab3 and Sen1 are not required for RNAPII termination.** To measure the global impact of mutations in NNS components on RNAPII transcription in *S. pombe*, we analyzed genome-wide RNAPII (Rpb1) occupancy by ChIP-seq using *nab3Δ, sen1Δ, and db18Δ* (Sen1 paralog) single mutants as well as a *sen1Δ db18Δ* double mutant. The absence of Nab3, Sen1, and Dbl8 did not result in noticeable read-through transcription at mRNA and snoRNA genes compared to the wild-type (WT) control. In contrast, a deficiency in the essential protein Seb1 results in defective termination at most RNAPII-transcribed genes (see *Pamt-seb1* mutant; Fig. 1a, b). Given that Nab3, Sen1, and Dbl8 do not copurify with Seb1, and considering the absence of termination defects in *nab3Δ, db18Δ, sen1Δ, and sen1Δ db18Δ* mutants (Fig. 1a, b), we conclude that NNS-dependent transcription termination is not conserved in fission yeast.

**Recruitment of mRNA 3’ end processing factors at ncRNA genes.** The lack of conservation of NNS-dependent transcription termination raises the question as to the mechanism of 3’ end processing and transcription termination at ncRNA genes in fission yeast. One clue about how ncRNAs could be processed in *S. pombe* derives from studies on the nuclear poly(A)-binding protein Pab2, the homolog of human PABPN1. We previously uncovered a polyadenylation-dependent pathway required for the maturation of independently transcribed snoRNAs that relies on Pab2. To test for the possibility that mRNA cleavage and polyadenylation factors function in 3’ end processing of ncRNA genes in *S. pombe*, we used ChIP-seq to examine the genome-wide-binding profile of proteins with conserved roles in mRNA 3’ end processing and transcription termination. This analysis included two independent components of the cleavage and polyadenylation factor I (CFI) complex, namely Rna14 and Pcf11, the endonuclease Ysh1 that is responsible for pre-mRNA cleavage before 3’ end polyadenylation, and Dhp1, the homolog of *S. cerevisiae* Rat1 and human XRN2 that promote transcription termination of RNAPII via 5’–3’ exonucleolytic activity. The ChIP-seq profile of Seb1 was also included, as this protein was shown to be important for 3’ end processing of both mRNAs and ncRNAs. Consistent with roles in 3’ end processing at mRNA
genes, Rna14, Pcf11, Ysh1, Seb1, and Dhp1 all showed strong binding at the 3′ end of protein-coding genes (Fig. 1c, f). Notably, they also exhibited robust enrichment at the 3′ end of independently transcribed snoRNAs (Fig. 1d, g) and snRNAs (Fig. 1e) genes. In contrast, these factors were not enriched at the 3′ end of intron-encoded snoRNAs (Supplementary Fig. 1), consistent with a maturation pathway independent of cleavage and polyadenylation. Interestingly, analysis of peak distribution based on average-binding profiles indicated that the recruitment of Dhp1 generally occurred downstream of the endonuclease Ysh1 (Fig. 1f, g; note that the peak of the green curves is shifted downstream relative to the pink curves). This is consistent with transcription termination of coding and ncRNA genes following the torpedo model, in which Dhp1 requires Ysh1-dependent endonucleolytic cleavage to promote disengagement of RNAPII from the DNA template5–7. On the basis of these results, we conclude that mRNA 3′ end processing factors are recruited at the 3′ end of both coding and ncRNA genes in fission yeast.

Synthesis of ncRNAs requires mRNA 3′ end processing factors. The presence of cleavage and polyadenylation factors at the 3′ end of ncRNA genes suggested a role for the mRNA 3′ end processing machinery in snoRNA synthesis. To examine the requirement of Rna14, Pcf11, Ysh1, and Dhp1 in the synthesis of ncRNAs, conditional strains were generated since these proteins

**Fig. 1** *S. pombe* mRNA 3′ end processing and transcription termination factors are recruited at the 3′ end of independently transcribed snoRNA and snRNA genes. **a, b** Average ChIP-seq profiles of total RNAPII (Rpb1) relative to mRNA poly(A) site (**a** n = 4755) and to annotated 3′ end of independently transcribed monocistronic snoRNAs (**b** n = 31) in WT (solid line, orange) and NNS mutant strains (dotted lines, orange) grown in rich medium (YES). Average ChIP-seq profiles of total RNAPII (Rpb1) in WT and Seb1-deficient cells (Pnmt-seb1) grown for 15 h in thiamine-supplemented minimal medium (EMM) are also shown (gold). Gene coordinates are available in Supplementary Data 1–2. **c–e** Normalized ChIP-seq signal of RNAPII (Rpb1) and the indicated mRNA 3′ end processing factors across the *fba1* mRNA (**c**), the *snR99* snoRNA (**d**), and the *snu5* snRNA (**e**) genes. **f, g** Average ChIP-seq profile of the indicated mRNA 3′ end processing factors over the same groups of mRNA (**f**) and snoRNA (**g**) genes as **a, b**.
are all encoded by essential genes. As used previously to repress the expression of essential core exosome subunits \textsuperscript{29} and Seb1 \textsuperscript{24}, we constructed strains with endogenous \textit{pcf11} and \textit{dhp1} genes under control of the thiamine-repressible \textit{nmt81} promoter (\textit{Pnmt}; see Methods for details). Importantly, levels of Seb1, \textit{Pcfl1}, and \textit{Dhp1} expressed from the \textit{nmt} promoter were similar to levels expressed from their endogenous promoter in non-repressing conditions, but showed robust depletion in thiamine-supplemented medium when expressed under the \textit{nmt} promoter (Supplementary Fig. 2a–c). Consistent with \textit{Pcfl1} and \textit{Dhp1} being essential for viability \textsuperscript{30}, \textit{Pnmt-pcf11} and \textit{Pnmt-dhp1} cells cultured in thiamine-supplemented medium showed growth arrest (Supplementary Fig. 2d). Depletion of \textit{Pcfl1} in \textit{S. pombe} also affects mRNA production \textsuperscript{31}, as expected of a protein essential for mRNA synthesis. To generate conditional strains for \textit{rna14} and \textit{ysh1}, we used the rapamycin-dependent anchor-away system \textsuperscript{32} (Supplementary Fig. 2e), as the \textit{nmt} promoter did not provide sufficient \textit{Rna14} and \textit{Ysh1} depletion to induce growth arrest. Importantly, inactivation of \textit{Ysh1} and \textit{Rna14} by nuclear depletion (Supplementary Fig. 2f, g) impaired mRNA synthesis in a rapamycin-dependent manner (Supplementary Fig. 2h, i) and resulted in the production of read-through transcripts (Supplementary Fig. 2j), consistent with 3′ end processing defects at protein-coding genes. To assess the contribution of \textit{Pcfl1}, \textit{Dhp1}, \textit{Ysh1}, and \textit{Rna14} to snoRNA 3′ end processing, we used an RNase H cleavage assay that can simultaneously detect both the mature snoRNA and 3′-extended polyadenylated snoRNA precursors (Fig. 2a). As a control, we used the \textit{pab2Δ} mutant that accumulates 3′-extended polyadenylated precursors, resulting in 45 and 34% reduction in \textit{snR3} and \textit{snR99} snoRNAs, respectively (Fig. 2b, c, compare lanes 3–4 to 1–2). Similarly, deficiencies in \textit{Pcfl1}, \textit{Ysh1}, and \textit{Rna14} resulted in reduced levels of mature \textit{snR3} and \textit{snR99} snoRNAs: 39 and 42% decrease, respectively, for \textit{Pcfl1} (Fig. 2b, c, lanes 5–6; Supplementary Fig. 2k); 45 and 46% decrease, respectively, for \textit{Ysh1} (Fig. 2d, e, lanes 3–4); and 41 and 33% decrease, respectively, for \textit{Rna14} (Fig. 2d, e, lanes 5–6). In addition to polyadenylated pre-snoRNA detected by RNase H assays (Fig. 2b, c), long \textit{snR3} read-through products were also detected by Northern blot analysis in cells deficient for \textit{Pcfl1}, \textit{Ysh1}, and \textit{Rna14} (Fig. 2f, lanes 2–3 and 5). In contrast, snoRNAs 3′-extensions did not accumulate in cells deficient for \textit{Dhp1} (Fig. 2b, c, lanes 9–10; Supplementary Fig. 2k; and Fig. 2f, lane 7), consistent with a role in transcription termination that occurs beyond 3′ end processing. As shown previously \textsuperscript{24}, a deficiency in Seb1 did not negatively impair RNA production as demonstrated by the similar levels of mature snoRNA between wild-type and Seb1-deficient cells, but changed cleavage site selection as observed by the lengthening of the 3′-extended snoRNA precursors on RNase

---

\textbf{Fig. 2} The mRNA cleavage and polyadenylation complex is required for snoRNA synthesis. \textbf{a} Schematic of the RNase H cleavage assay used in \textbf{b}–\textbf{e}. After \textit{RNase H} cleavage of the snoRNA at sites of RNA:DNA hybrids, the 3′ fragment (mature or 3′-extended) is detected by Northern blotting (NB). Addition of oligo (T) to the RNase H reaction removes heterogeneous poly(A) tails, generating discrete products. \textbf{b}, \textbf{c} Total RNA prepared from the indicated strains that were previously grown in thiamine-supplemented medium for 15 h to deplete Pcf11, Seb1, and Dhp1 was treated with RNase H in the presence of DNA oligonucleotides complementary to \textit{snR3} (\textbf{b}) and \textit{snR99} (\textbf{c}). RNase H reactions were performed in the presence (+) or absence (−) of oligo (T). The top panel represents a longer exposure of the middle panel to see 3′-extended (3′-ext) cleavage products. The 5S rRNA was used as a loading control. \textbf{d} Northern blot analysis using total RNA prepared from the indicated strains that were treated with either rapamycin (lanes 1–3) or thiamine (lanes 4–6). The blot was hybridized using DNA probes specific to \textit{snR3} and the 18S rRNA. The position of mature \textit{snR3}, 18S and 25S rRNAs, as well as 3′-ext \textit{snR3} read-through (RT) products is indicated on the right.
Widespread polyadenylation at snoRNA genes. To obtain a comprehensive view of pre-snoRNA polyadenylation and to address whether differences exist between 3′ end processing of H/A/ACA and C/D box snoRNAs, we used 3′ region extraction and deep sequencing (3′READS), an approach developed to map mRNA polyadenylation sites at the genome-wide level. Over 6 millions reads mapping to polyadenylation sites (PAS reads) were obtained from wild-type S. pombe cells grown in rich and minimal media by 3′READS. Although the majority (~60%) of unique PAS reads mapped to 3′ UTR of protein-coding genes, roughly 18% of unique PAS locations mapped to intergenic regions. Notably, out of 31 independently transcribed monocistronic snoRNAs, 24 (77%) and 27 (87%) had a mappable poly(A) site located downstream of a snoRNA in minimal and rich media, respectively (Supplementary Data 2). It should be noted that snoRNA genes for which PAS reads could not be identified showed barely detectable levels of RNAPII as determined by ChIP-seq assays. Comparison between C/D and H/A/ACA box snoRNAs indicated similar proportions of polyadenylated species (Fig. 3a). Figure 3b shows 3′READS data for the H/A/ACA box snR99 snoRNA. As can be seen, 3′READS identified a major PAS located 78 nt downstream of the annotated snR99 mature 3′ end in wild-type cells. In the pab2Δ mutant, 3′READS detected a ~45-fold increase in the abundance of snR99 precursors that used this major PAS (Fig. 3b). Notably, whereas depletion of Seb1 resulted in the use of a single distal PAS located 118 nt downstream from the major snR99 PAS (196 nt from the snR99 mature 3′ end), the two different
PASs identified in the WT, pab2Δ, and Pmnt-seb1 strains were used in cells deficient for Pcf11 (Fig. 3b). Collectively, the 3′READS data are entirely consistent with results obtained using RNase H cleavage assays (see Fig. 2c).

Global analysis of 3′READS data revealed that the median distance between the annotated (mature) snoRNA 3′ end and the pre-snoRNA cleavage site is 77-nt in wild-type cells grown in rich medium (Fig. 3c, red boxes). Consistent with previous results24 (Fig. 2), the absence of Pab2 did not generally affect cleavage site selection at snoRNA genes (Fig. 3c, green boxes), but resulted in a significant accumulation of polyadenylated snoRNA precursors (Fig. 3d, compare green and red boxes); interestingly, HACA box pre-snoRNAs appeared to accumulate to greater levels as compared to C/D box precursors in the absence of Pab2 (Fig. 3d). In contrast, the absence of Pab2 did not generally impact the levels of polyadenylated mRNAs (Supplementary Fig. 3). In the case of Seb1, as seen for snr99 (Fig. 3b), we noted a clear shift of the general distribution of polyadenylated pre-snoRNAs toward longer 3′ extensions (Fig. 3c, blue boxes); the median distance between the snoRNA mature 3′ end and its corresponding poly(A) site increased from 79-nt in wild-type cells grown in minimal medium to 131-nt in Seb1-deficient cells (compare blue and yellow boxes). This result is consistent with the preferential use of distal cleavage sites in the seb1 mutant (Fig. 2), an outcome also observed for mRNAs34 (Supplementary Fig. 3). The overall distribution of pre-snoRNA polyadenylation sites was also extended in Pcf11-deficient cells, a result that was greater for HACA box snoRNAs (Fig. 3c, compare purple and yellow boxes), as well as for mRNAs34. In addition, a general accumulation of polyadenylated pre-snoRNAs was observed in pcf11 and seb1 mutants (Fig. 3d, compare purple and blue boxes to yellow boxes). In summary, most independently transcribed snoRNA genes produce 3′ extended polyadenylated precursors in S. pombe.

A mechanism of snoRNA 3′ end processing that involves cleavage and polyadenylation by mRNA maturation factors posits that cis-acting elements present in the 3′ flanking region of coding and ncRNA genes should be similar. We and others have shown that poly(A) signals around S. pombe mRNAs are associated with upstream AAU[A/G]AA hexamers and downstream GU A motifs34–36. However, the relatively small number of independently transcribed snoRNAs (n = 31), as compared to mRNAs (n = 4755), limited the identification of significantly enriched sequence motifs around snoRNA poly(A) sites. We therefore tested whether cis elements promoting mRNA 3′ end processing could support 3′ end maturation of a snoRNA. For this, we generated constructs in which 1-kb of snr99 downstream sequence was exchanged with 1-kb of downstream sequence from rps2 and trxl protein-coding genes (Fig. 3e). These three constructs as well as a vector control were chromosomally integrated as a single copy into a snr99-null strain. As shown in Fig. 3f, 3′ flanking sequences from both rps2 and trxl supported proper 3′ end formation and normal snr99 accumulation, as demonstrated by the length and expression level of the snoRNA, which were similar to snr99 expressed from its authentic downstream sequence (compare lanes 3–5). Furthermore, RNase H cleavage assays and 3′ RACE experiments revealed Pab2-dependent polyadenylated snoRNA precursors that used the natural rps2 and trxl cleavage sites (Fig. 3g and Supplementary Fig. 3c, d). Based on these findings, we conclude that snoRNA 3′ end processing in S. pombe occurs via a pathway analogous to mRNA biogenesis.

Ty1-P and Ser2-P peak at the end of mRNA and ncRNA genes. A number of studies have examined the association between the status of CTD phosphorylation and 3′ end processing/transcription termination. Since little is known about the association between CTD modifications and RNA processing in fission yeast, we mapped Tyr1, Ser2, Ser5, and Ser7 phosphorylation at the genome-wide level in wild-type S. pombe. ChiP-seq data from independent biological replicates showed high correlation throughout the S. pombe genome (see Supplementary Fig. 4). Averaging ChiP-seq signal across protein-coding genes revealed CTD phosphorylation patterns similar to those characterized in S. cerevisiae (Fig. 4a): Ser5 and Ser7 phosphorylation (Ser5-P and Ser7-P) marks peaking at the 3′ end of coding genes after the transcription start site (TSS), whereas Ser2 phosphorylation (Ser2-P) steadily increased along the gene body and sharply peaked downstream of the poly(A) site. The pattern of Ty1 CTD phosphorylation (Ty1-P) at protein-coding genes was similar to that of Ser2-P, peaking sharply downstream of the poly(A) site (Fig. 4a). It is interesting to note that the genome-wide pattern of Ty1-P signal in S. pombe contrasts to that of S. cerevisiae in which Ty1-P signal was shown to peak upstream of the poly(A) site20. Occurrence profiles of CTD phosphorylation marks at snoRNA genes showed a pattern somewhat similar to that of mRNA genes, with peaking Ser5-P and Ser7-P signals at the 5′ end that started decreasing in the snoRNA gene body (Fig. 4b). As seen for mRNA genes, Ser2-P and Ty1-P signal steadily increased along the body of snoRNA genes, with peaks that coincided with the decline in total RNAPII downstream of snoRNA poly(A) sites (Fig. 4b). Together, our findings indicate that termination of RNAPII transcription at mRNA and snoRNA genes in S. pombe occurs at regions where maximal signal of Ser2 and Ty1 phosphorylation are detected.

Pcf11 and Seb1 have CTD-interacting domains (CID) that preferentially recognize Ser2-phosphorylated RNAPII in vitro19,25. Therefore, we compared the ChiP-seq profiles of Pcf11 and Seb1 to genome-wide RNAPII Ser2-P signal. As shown for a protein-coding and a snoRNA gene (Fig. 4c, d), the peak occupancy of Pcf11 and Seb1 binding generally coincided with regions of maximum Ser2-P signal that were found downstream of poly(A) sites. The average occupancy profiles of Pcf11 and Seb1 also colocalized with regions of maximal Ty1-P signal (Fig. 4c–f). Globally, ChiP-seq signals of 3′ end processing factors were found to be more correlated genome-wide with Ser2 and Ty1 phosphorylation than with Ser5 and Ser7 phosphorylation (Figs. 4g, 0.58 vs 0.25). These results support the view that Ser2 and Ty1 CTD phosphorylation are functionally relevant to fission yeast 3′ end processing and transcription termination.
RNAPII component (Rpb3-HA; Fig. 5a–c). Importantly, transcription termination defects are a general feature of Dhp1-deficient cells, as seen by a clear shift in the Rpb1 and Rpb3-HA average signals downstream of the noticeable decline observed in wild-type cells (Fig. 5d, e; compare solid and dotted lines). These results show that Dhp1 is required for RNAPII termination at both coding and ncRNA genes.

Hyperphosphorylation of CTD repeats on Ser2 was previously observed in a temperature-sensitive mutant of S. cerevisiae rat37. To address how Dhp1 inactivation globally affects CTD phosphorylation at transcribed S. pombe genes, we compared Tyr1, Ser2, Ser5, and Ser7 phosphorylation levels between Dhp1-deficient and control cells using spike-in-normalized ChIP-seq. As shown in Fig. 5f, g, inactivation of Dhp1 resulted in a generalized increase in Ser2-P and Tyr1-P signal, which was most noticeable at the 3′ end of genes. Analysis of Ser2-P/total RNAPII ratios at the 3′ end of genes suggests that the levels of Ser2 phosphorylation does not truly increase in Dhp1-deficient cells, but that the CTD remains phosphorylated on Ser2 for an extended period of time after cleavage (see Supplementary

**Fig. 4** Tyr1-P and Ser2-P forms of the RNAPII CTD colocalize with 3′ end processing factors at coding and ncRNA genes. a, b Average ChIP-seq profile of total RNAPII (Rpb1) and the indicated CTD modifications in a WT strain across 4755 mRNAs (a) and 24 monocistronic snoRNAs (b) with a mapped poly (A) site in minimal medium. c, d PAS read density (3′ READS) and normalized ChIP-seq signal of total RNAPII (Rpb1), the indicated CTD modifications, as well as Pcf11 and Seb1 across the fba1 mRNA (c) and the snR99 snoRNA (d) genes. e, f Average profile of total RNAPII (Rpb1), Ser2-P, Tyr1-P, Pcf11, and Seb1 relative to the poly(A) site of 4755 mRNA (e) and 24 monocistronic snoRNA (f) genes in minimal medium. g Genome-wide pairwise Pearson correlation coefficient matrix at a resolution of 10 bp followed by a hierarchical clustering. Rectangles include the average of a subset of correlation values between 3′ processing factors and CTD modifications.
snoRNA transcription termination requires Ysh1 cleavage. The global transcription termination defects at mRNA and snoRNA genes in Dhp1-deficient cells suggest that coding and snoRNA genes are terminated by torpedo in fission yeast. This mechanism entails an endonucleolytic cleavage prior to Dhp1-dependent transcription termination. To test whether the endoribonuclease subunit of the cleavage and polyadenylation specificity factor (CPSF) complex, Ysh1 (CPSF-73 in humans), is required for transcription termination of snoRNA genes in fission yeast, we analyzed RNAPII density in cells where Ysh1 nuclear depletion was achieved by galactose induction of the GAL1-10 promoter. As expected, inactivation of Ysh1 resulted in a clear read-through phenotype at a protein-coding gene (Fig. 6a). Globally, the RNAPII 3′ peak showed a clear shift downstream in the ysh1 mutant as compared to control cells (Fig. 6d). Importantly, the consequences of Ysh1 nuclear depletion were comparable at the snr99 gene (Fig. 6b), causing a marked downstream shift in the average RNAPII profile at snoRNA genes (Fig. 6c). pcf11 and rna14 mutants also showed clear evidence of read-through transcription at snr99 and snr3 genes (Supplementary Fig. 6), consistent with a general role of the CPF in transcription termination at snoRNA genes in fission yeast. The transcription termination defects observed at snoRNA genes in the CPF

Fig. 5b). In contrast, Tyr1-P/total RNAPII ratios showed an increase in the maximal signal of Tyr1 phosphorylation in the dhp1 mutant (Supplementary Fig. 5b), suggesting that more CTD repeats have a phosphorylation mark on Tyr1 in Dhp1-depleted cells compared to wild-type cells. Prolonged Pcf11 binding at the 3′ end of coding (Supplementary Fig. 5c–f) and 24 monocistronic snoRNA (e, g) genes with a mapped p(A) site in thiamine-treated minimal medium (d–f) and 24 monocistronic snoRNA (e, g) genes with a mapped p(A) site in thiamine-treated minimal medium.
mutants (ysh1, pcf11, and rna14) are thus consistent with the detection of long read-through transcripts (Fig. 2f), which are not detected in the pab2Δ mutant (Fig. 2f, lane 8) that does not show transcription termination defects.28 Interestingly, whereas Dhp1 was required for efficient transcription termination of both snoRNA and snRNA genes (Fig. 5b, c and Supplementary Fig. 5a), the nuclear depletion of Ysh1 did not result in read-through transcription at snRNA genes (Fig. 6c and Supplementary Fig. 7a).

We next addressed whether the endonucleolytic activity of Ysh1 was required to promote termination of snoRNA transcription. We generated ysh1 alleles expressing single amino acid substitutions at conserved histidine residues that were shown to contact critical zinc atoms in the active site of human CPSF.39 We used a “complementation after nuclear depletion” approach by chromosomal integration of the different ysh1 alleles into the ysh1 anchor-away strain and confirmed that the corresponding proteins were expressed (Supplementary Fig. 7b). After nuclear
The endonucleolytic activity of Ysh1 is necessary for termination of snoRNA transcription. a–c Normalized ChIP-seq signal of total RNAPII (Rpb1) in WT (top) and ysh1 mutant (bottom) cells across the fba1 mRNA (a), the snR99 snoRNA (b), and the snu5 snoRNA (c) genes grown in rapamycin-treated minimal medium for 4 h. The dashed-line rectangles highlight transcriptional read-through at fba1 and snR99 genes in Ysh1-deficient cells. d, e Average ChIP-seq profile of RNAPII (Rpb1) in WT (solid lines) and Ysh1-deficient (dotted lines) cells across 4755 mRNA (d) and 24 monocistronic snoRNA (e) genes grown in rapamycin-treated minimal medium for 4 h. f, g RNAPII ChIP-qPCR analysis on the fba1 (f) and snR99 (g) genes using extracts prepared from either wild-type (WT) or ysh1 anchor-away (ysh1-AA) strains containing chromosomally integrated constructs that express the indicated versions of FLAG-tagged Ysh1 (WT, H403F, and H165F) as well as an empty vector (EV) control. Bars above the fba1 and snR99 genes show the positions of PCR products used for ChIP-qPCR analyses. Cells were grown in the presence of rapamycin for 4 h to deplete endogenous Ysh1 from the nucleus. ChIP signals (percent of input) were normalized to region 1. Error bars indicate SD. n = 3 biological replicates from independent cultures. h–j Northern blot analysis of fba1 (h), snR3 (i), and snR99 (j) genes using total RNA prepared from using total RNA prepared from the same WT (lane 1) and ysh1 anchor-away (ysh1-AA, lanes 2–5) strains as in f, g. The position of read-through (RT) transcripts is indicated on the right. The 18S rRNA was used as a loading control.

Depletion of endogenous Ysh1 by rapamycin, the expression of catalytic mutant versions of Ysh1 resulted in growth arrest (Supplementary Fig. 7c), consistent with the catalytic activity of Ysh1 being required for cell viability.41 Next, we used RNAPII ChIP-qPCR assays to examine the extent to which the mutant versions of Ysh1 restored the transcription termination defects induced by the nuclear depletion of endogenous Ysh1. As a control, expression of wild-type Ysh1 in the ysh1 anchor-away strain prevented the reappearance of transcriptional read-through in RNAPII levels downstream of both protein- and snoRNA-coding genes (Fig. 6f, g. compare dark blue, gray, and orange bars). In contrast, the endonuclease mutant versions of Ysh1 showed increased in RNAPII density at the 3′ end of both mRNA and snoRNA genes (Fig. 6f, g. compare yellow and light blue bars to gray bars). Consistently with these RNAPII ChIP-qPCR data, RNA analysis showed the accumulation of read-through transcripts in cells that expressed mutant versions of Ysh1 (Fig. 6h–j). compare lanes 4–5 to lane 3; Supplementary Fig. 7d). We conclude that the endonucleolytic activity of Ysh1 is required for 3′ end processing and transcription termination of both mRNA and snoRNA genes.

Transcription termination at long noncoding RNA genes. Long noncoding RNAs (lncRNAs) contribute a substantial and diverse portion of noncoding transcriptions. In S. pombe, over 85% of the roughly 1500 annotated lncRNAs are expressed below one copy per cell during the mitotic cell cycle,42 thus making it challenging to detect lncRNA expression under normal conditions. Furthermore, a large proportion of the S. pombe lncRNAs are antisense to protein-coding genes, rendering analysis of non-stranded ChIP-seq data unreliable. Attempts to obtain a global portrait of CTD phosphorylation patterns and CPF binding at lncRNA genes in fission yeast were unsuccessful, presumably due to their low levels of transcription. We therefore focused our attention on a handful of lncRNA genes that showed robust levels of RNAPII during vegetative growth as determined by our ChIP-seq analysis of Rpb1 and Rpb3-HA, and observed clear enrichment of Ser2-P and Tyr1-P together with binding of CPF and termination factors at the 3′ end of the SPNCRNA.532 (Fig. 7a) and SPNCRNA.491 (Fig. 7b) lncRNA genes. Importantly, read-through transcription was detected downstream of SPNCRNA.532 (Fig. 7a) and SPNCRNA.491 (Fig. 7b) in cells deficient for Pcf11, Rna14, Ysh1, and Dhp1 (Fig. 7a, b and Supplementary Fig. 6c, d), consistent with a cleavage-dependent mechanism of transcription termination. Similar binding profiles were found for a recently described lncRNA, SPNCRNA.1459/nam1 (Fig. 7c and Supplementary Fig. 6e), which represses entry into the meiotic differentiation cycle.43 Although additional work will be required to get a global view of RNA processing and CTD phosphorylation at S. pombe lncRNA genes, these data suggest a mechanism of transcription termination that relies on endonucleolytic cleavage and Dhp1-dependent torpedo.

Discussion

Despite extensive characterization of how the NNS complex promotes termination at ncRNA genes in S. cerevisiae, whether NNS-like transcription termination is conserved across eukaryotic species was unclear. Despite the presence of homologs for all of the NNS components, our results indicate that NNS-mediated termination is not conserved in the distantly related yeast S. pombe. In the absence of an NNS-like complex, our findings reveal that a universal cleavage-dependent mechanism that involves the conserved Dhp1 5′–3′ torpedo nuclease is used to terminate transcription of both mRNA and ncRNA genes.

Two key indications support the absence of an NNS-like transcription termination pathway in fission yeast. First, Seb1 (Nrd1 homolog), Nab3, Sen1, and Dbl8 do not form a stable complex in fission yeast.24,25,44 Second, transcription termination defects were not detected in nab3Δ, sen1Δ, dbl8Δ, and sen1Δ dbl8Δ mutants. These data suggest that while NNS components have been conserved between budding and fission yeasts, they have acquired species-specific functions. For instance, S. pombe Seb1 recognizes Ser2-phosphorylated RNAPII and is recruited to the 3′ end of both mRNA and ncRNA genes.24,25 In contrast, S. cerevisiae Nrd1 preferentially binds to Ser5-phosphorylated RNAPII and is enriched at noncoding transcription units.2 Also consistent with the idea that NNS components have functionally diverged between the evolution of budding and fission yeasts from a common ancestor, Sen1 is primarily associated with RNA polymerase III (RNAPIII)-transcribed genes in S. pombe.44 An NNS-like transcription termination pathway is also unlikely to exist in metazoa, as clear homologs or functional analogs of Nrd1 and Nab3 components have not been described. Conversely, a homolog for the yeast Sen1 helicase has been characterized in humans (Senataxin, SETX), and a role for SETX in promoting efficient transcription termination at model protein-coding genes has been described.45,46 However, how SETX globally contributes to transcription termination of mRNA genes remains to be demonstrated, as data presented here as well as in S. cerevisiae47 indicate that yeast Sen1 is not required for efficient transcription termination at protein-coding genes.

By combining genome-wide studies and functional analyses of the fission yeast mRNA 3′ end processing machinery, we provide evidence supporting that a general cleavage-dependent, torpedo-mediated mechanism is used to promote transcription termination of protein-coding and ncRNA genes. Several observations support this conclusion: (i) mRNA 3′ end processing (Rna14, Pcf11, and Ysh1) and termination (Dhp1) factors are specifically recruited at the 3′ end of independently transcribed snoRNA, snRNA, lncRNA genes; (ii) mRNA 3′ end processing factors are required for snoRNA synthesis; (iii) polyadenylated pre-snoRNAs are produced from most snoRNA genes; (iv) 3′ end cis-sequences of protein-coding genes can promote snoRNA 3′ end maturation and accumulation; and (v) transcription termination at ncRNA
genes requires endonucleolytic cleavage and the torpedo nuclease Dhp1. These data support a model in which cleavage and polyadenylation are essential steps for 3' end processing and transcription termination at both mRNA and snoRNA genes. Although it has been shown that particular components of the mRNA 3' end processing machinery, such as Pcf11, Rat1, and Glc7, are recruited to short ncRNA genes in *S. cerevisiae*16,48–50, these factors appear to function either separately from the CPF complex to contribute to NNS-dependent termination or as a failsafe mechanism of transcription termination. This contrasts our findings in fission yeast, where the prevailing mechanism of transcription termination at ncRNA genes involves the cleavage and polyadenylation complex. Given the unexpected similarity of cleavage-dependent transcription termination at mRNA and snoRNA genes, it will be interesting to understand the mechanism used to distinguish between mRNAs and snoRNAs; whereas the mature mRNA needs to retain its poly(A) tail, the polyadenylated 3'-extended pre-snoRNA is targeted by Pab2-dependent exosome-mediated 3' end trimming to yield a mature non-polyadenylated snoRNA. We speculate that co-transcriptional decoration of nascent pre-mRNAs with nuclear export factors prevents nuclear retention and Pab2-dependent 3' end trimming, whereas the nucleolar-targeting signals of snoRNA ribonucleoproteins (snoRNPs) assembled during snoRNA transcription51 would promote nuclear retention and 3' end maturation.

**Fig. 7** Transcription termination at lncRNA genes and general model for 3' end processing and transcription termination of mRNA and snoRNA genes in fission yeast. a–c Normalized ChIP-seq signal of total RNAPII (Rpb1 and Rbp3-HA), the indicated CTD modifications, and 3' end processing factors across the SPNCRNA.532 (a), the SPNCRNA.491 (b), and the nam1/SPNCRNA.1459 (c) genes. d, e Recruitment of cleavage and polyadenylation factors (CPF) by poly(A) signals (PAS) is a common feature of mRNA and snoRNA genes (see Cleavage). Endonucleolytic cleavage by the CPF-associated Ysh1 nuclease will generate an RNAPII-bound unprotected 5' end that will be targeted by the 5'-3' exonuclease Dhp1, contributing to termination of mRNA and snoRNA transcription (see Termination). The RNA-bound CPF complex promotes snoRNA 3' end maturation, possibly by facilitating Pab2 recruitment to snoRNA precursors. The co-transcriptional recruitment of specific export factors to nascent mRNAs (d) may represent a decisive step that prevents Pab2-dependent exosome-mediated RNA processing in the nucleus (e).
In addition to the detection of read-through transcription at snoRNA genes, depletion of CPF components including Pcf11, Ysh1, and Rna14 also led to the accumulation of polyadenylated snoRNA precursors similar to the pad2A mutant. Thus far, our studies on 3’ end processing and transcription termination at snoRNA genes have identified three types of mutants: (i) one type showing transcription termination defects but no detectable RNA-processing deficiencies (dpfl1), (ii) one type showing RNA maturation defects but normal transcription termination (dpfl2), and (iii) mutants that show defects in both transcription termination and snoRNA 3’ end processing (pcf11, ysh1, rna14, and seb1). Collectively, these findings suggest that in addition to promoting the endonucleolytic cleavage for Dhp1-dependent termination, the CPF also contributes to snoRNA 3’ end processing in fission yeast. Although the precise role of the CPF in snoRNA 3’ end maturation remains to be determined, it may be important for the proper recruitment of Pab2, a view that is supported by protein interaction studies.

As for snoRNA genes, termination of snRNA transcription in S. cerevisiae appears primarily dependent on NNS. We found that where transcription termination at both snoRNA and snRNA genes required Dhp1 in S. pombe, termination of snRNA transcription was not affected in Ysh1-deficient cells. This result suggests that the 5’ entry point for Dhp1 loading at snRNA genes may be mediated independently of the nuclelease activity of the mRNA 3’ end processing machinery. Interestingly, read-through transcription at the S. pombe U2 snoRNA gene was previously reported in a temperature-sensitive mutant of pac1.35, which encodes a homolog of the S. cerevisiae RNase III-like endonuclease Rnt1. Although the identity of the nucleases responsible for cleavage at S. pombe snoRNA genes remains to be determined, our data suggest that snoRNA transcription termination in fission yeast is more closely related to humans than to S. cerevisiae, as mRNA 3’ end processing factors and the torpedo nuclease XRN2 are important for efficient transcription termination of human snoRNA genes.

Our study also elicits interesting questions as to the functional significance of RNAPII CTD phosphorylation in fission yeast. First, given the reciprocal relationship between Ser2 phosphorylation and 3’ end processing in S. cerevisiae and humans,44 the peak of Ser2-P signal at the 3’ end of both mRNA and ncRNA genes provides an additional line of evidence supporting co-transcriptional 3’ end processing by the CPF in fission yeast. Intriguingly, however, Ser2-P does not appear to be essential in S. pombe, suggesting that protein-RNA interactions may function as a redundant mechanism of CPF recruitment, as recently shown for elongation factors in budding yeast.59 Second, our data indicates that the pattern of Tyr1-P on the CTD of S. pombe RNAPII at the 3’ end of genes differs from the distribution of Tyr1-P in S. cerevisiae.30 Accordingly, although Tyr1-P levels increase along coding regions in both budding and fission yeasts, Tyr1-P signal decreases upstream of the poly(A) site in S. cerevisiae, whereas Tyr1-P signal persists downstream of the poly(A) site in S. pombe. These differing patterns of Tyr1-P around the poly(A) were unexpected, as Tyr1-P in S. cerevisiae was previously proposed to inhibit binding of CID-containing termination factors (Pcf11 and Rtt103) to Ser2-P CTD and prevent premature termination in coding regions,44, a role that would also be anticipated to contribute to general transcription in S. pombe. Yet, the similar distribution of Ser2-P and Tyr1-P profiles downstream of the poly(A) site in S. pombe argues for diverging roles of Tyr1-P between budding and fission yeasts. Consistent with this idea, a truncated version of the CTD in which a phe-nylalanine replaces Tyr1 (Y1F) in each heptad repeat results in lethality in S. cerevisiae,68 whereas the analogous truncated Y1F CTD mutant is viable in S. pombe.36

Our genome-wide analysis of RNAPII occupancy revealed that Dhp1 is necessary for transcription termination of most protein-coding genes in fission yeast. In addition to transcription termination defects, elevated levels of Ser2-P signal were found at model protein-coding genes in a temperature-sensitive mutant of S. cerevisiae rad1.77 Our data extend these findings by showing that (i) this observation is conserved in fission yeast, (ii) that Dhp1 depletion also causes elevated levels of Tyr1 phosphorylation, and (iii) that Dhp1 deficiency affects Tyr1 and Ser2 CTD phosphorylation at the genome-wide level. Although further studies will be needed to investigate how Dhp1/Rat1 nucleases modulate CTD phosphorylation at the 3’ end of genes, which may involve competition with the recruitment of CTD kinases,61 this finding highlights the important functional relationship between CTD phosphorylation and transcription termination.

Given that NNS-like transcription termination does not appear to be conserved in mammalian cells, it remains unclear how independently transcribed human snoRNAs, such as U3, U8, U13, and the human telomerase RNA, acquire their 3’ ends. The fact that we have recently uncovered a polyadenylation-dependent 3’ end maturation pathway for human telomerase RNA that depends on the nuclear poly(A)-binding protein PABPN1 supports the existence of an evolutionarily conserved role for Pab2/PABPN1 in the maturation of independently transcribed snoRNAs. It will be interesting to determine whether 3’ end processing and transcription termination of human snoRNA genes that are characterized by independent transcription unit depend on the mRNA cleavage and polyadenylation complex, as unveiled by our findings using fission yeast.

Methods

Yeast strains and media. A list of all S. pombe strains used in this study is provided in Supplementary Table 1. Cells were routinely grown at 30 °C in Edinburgh minimal media (EMM) or in yeast extract medium (YES) supplemented with adenine, uracil, histidine, and leucine. Conditional strains, in which the genomic copy of the essential genes pcf11, dhp1, and seb1 are expressed from the thiamine-sensitive nmt41 and nmt81 promoters (Supplementary Table 1), were repressed by the addition of 60 µM of thiamine in the EMM media for 12–15 h as previously described.12,49 Depletion of nuclear Ysh1 and Rna14 were done by an anchor-away strategy. Briefly, wild-type (WT), Ysh1-FRB-GFP (ysh1-1A), and Rna14-FRB-GFP (rna14-1A) strains were grown in either EMM without leucine or Yeast extract, and rapamycin was added at a final concentration of 2.5 µg/ml for 4–15 h. Cells were collected at OD600nm of 0.5–0.8. Gene disruptions and C-terminal tagging of proteins were performed by PCR-mediated gene targeting using lithium acetate method for yeast transformation. Expression of tagged proteins was confirmed by western blotting and knockouts strains by the absence of RNA by RT-PCR.

Growth assays. Exponential cultures of S. pombe cells were adjusted to an OD600nm of 1.0, and serially diluted tenfold using water. Each dilution was spotted (3 µl per spot) on EMM plates with or without 15 µM of thiamine in the case of strains using the thiamine-sensitive nmt41 promoter (Pnmt-), or on YES plates containing DMSO or 2.5 µg/ml rapamycin for the analysis of anchor-away mutants. Plates were incubated at 30 °C for 2–7 days.

Microscopy. Ysh1-GFP and Rna14-GFP localization was detected by fluorescence microscopy. Briefly, precultures grown in EMM supplemented with adenine, uracil and histidine (EMM Leu−) were used to inoculate larger 25-ml cultures that we grown in EMM leu− to early log phase (OD600nm 0.25–0.3). Rapamycin was then added to a final concentration of 2.5 µg/ml (same volume of DMSO was added in parallel as control) and samples were taken at 0, 0.5, 1, 2, 3, 4, and 24 h. Cells were diluted 1:10 in water. Nuclei were stained using Hoechst 33342 for 5 min (0.2 mg/ml) and live cells were mounted on a 1.2% agarose/EMM leu− pad as described previously.44 GFP-tagged proteins and nuclei were detected at 470 nm and 365 nm, respectively, using a Colibri system (Carl Zeiss Canada, Toronto, ON, Canada) on a Zeiss Axio Observer Z1 inverted microscope with a x60/1.4 oil objective. Data were analyzed using the ZEN black software (Carl Zeiss Canada).

RNA preparation and analyses. Total RNA was extracted using the hot-acid phenol method. RNA samples were resolved on agarose-formaldehyde gel or on 6% polyacrylamide-8M urea gels, transferred onto nylon membranes and probed as
described previously24. RNase H assays were performed as described previously with 15 µg of RNA25,26. Briefly, total RNA was RNAse H-treated in a mixture containing RNase H and complementary DNA for 1 h and release a specific 3' fragment, as well as with or without oligo (d)T to remove the heterogeneity of poly(A) tails. Preparation of probes, blot hybridization, washes, and quantification of signals were done as described previously25,26. Uncovered blots are shown in Supplementary Fig. 9.

Chromatin immunoprecipitation (ChIP) assays. ChIP-qPCR and ChIP-seq experiments were performed as described previously24. Chromatin was immunoprecipitated directly using Pan Mouse IgG Dynabeads (Life Technologies, 11040) for TAP-tagged proteins. ChIP were also performed using Pan Mouse IgG Dynabeads coated with antibodies specific to Rpb1 (clone 8WG16; Covance, MMS-126R) or the hemagglutinin (HA) sequence (clone 12CA5, Roche, 1158381601). ChIP with phospho-CTD antibodies were performed using 2 µg of anti-CTD antibody: Ser2-P (Abcam, ab5095) and Ser5-P (Abcam, ab1311) were coated to Dynabeads M-280 Sheep anti-rabbit IgG beads (Life Technologies, 11203D), respectively. All experiments were performed as described previously24. Chromatin was immunoprecipitated using antibody-coated beads. ChIP-seq experiments were prepared as described previously24 using either the NEBNext ChIP-seq Library Prep Master Mix Set for Illumina kit (New England Biolabs) or the SPARK DNA Sample Prep Kit Illumina Platform (Enzymatics) according to the manufacturer's instructions.

ChIP-seq processing. Briefly, the raw reads were trimmed using Trimmomatic version 0.3225 with the algorithm mean or min on the read length (above or below 70 nt) and default parameters, onto a concatenated reference genome containing both the sequences of the S. pombe ASM294v2 and S. cerevisiae S288C assemblies. Normalization and other analyses were performed by Orlando et al.38. Notably, the signal mapping was performed in order to avoid discarding the signal at regions of the genome that are duplicated (BWA is randomly assigning the reads), but we generated mapability tracks for various read length to help the interpretation of particular regions. The number of reads mapped on the concatenated genomes varied from 91.5 to 99.6% of the sequenced reads (Supplementary Data 3). For each data set, the sum of the signal coming from the reads aligned to the S. cerevisiae genome (Supplementary Data 3). Signal density files in BedGraph format were then generated using BEDTools genomcov version 2.1727 with default parameters, then converted in uniform 10 nt bins Wig files for further normalization steps (inspired by the script bedgraph_to_wig.py). The wig tracks were submitted to the UCSC Genome Browser using Orlando et al.38. Mouse and rat were associated to the closest gene, including ncRNAs. The GTF genome annotations of S. pombe (Spagnol and expression were confirmed by RT-qPCR using primers located in the 5'UTR sequences. The signal density files were then encoded in bigWig format using the Kent utilities visualization of the data were performed using AssemblyHub on the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgHubConnect?hubURL=https://databhub-8kims5w.ute.ugapen.ca/CTD_sno_pombe/hub&hghubHub.do_firstDb=on&hghubHub.do_redirect=on&hghubConnect.removeTrackHub=on).

3' READS analysis. As described in Liu et al.34,35, 3' READS reads were mapped onto the S. pombe ASM294v2 assembly using Bowtie 2 (local mode)36. Uniquely mapped reads (with MAPQ score >10) that had at least two additional 5' Ts after genome alignment were considered as polyA sites (PAS) reads. Each PAS was then assigned to the closest gene, including ncRNAs. The GTF genome annotations file version ASM294v2.29 was downloaded from PomBase125,126 in May 2017. However, based on manual inspection of our ChIP data, the orientation of three snoRNAs (namely, SPNSNORNA.36, SPNSNORNA.37, and SPNSNORNA.41) was inverted. In more details and based on Liu et al.34,35, PAS supported with at least two reads outside CDS were associated to the closest 3' end gene annotation in a strand-specific manner to avoid cross-platforms annotations in the same region. The software nbGnomes was used to generate the signal for each gene. The genome was split into 500 bp regions and the definition of the closest 3' end was inferred by the number of reads in each region. For each gene, the strongest PAS was positioned to the closest gene, including ncRNAs. The GTF genome annotations file version 1.0.0 was used to generate the average coverage with the following common parameters: annotation mode, absolute analysis method, 10 bp windows size, mean coverage.

Plasmid pFB1337. To generate amino acid substitutions at residues involved in Ysh1 endonucleolytic activity, histidine (H) residues 165 or 403 of Ysh1 were substituted by cysteine (C) to create pFB1335 (H165F) and pFB1338 (H403F), respectively. All mutations were performed by site-directed mutagenesis using pFB1337 as a template. All DNA constructs were confirmed by DNA sequencing. For single integration into the ade6 locus, pFB866, pFB1337, pFB1335, and pFB1338 were linearized and transformed into FY2066 and/or FY2110 strains. Positive integrates were confirmed by growth selection on EMG agar plates lacking adenine and leucine and by western blotting.

Library preparation and Illumina sequencing. DNA libraries for ChIP-seq experiments were prepared as described previously14 using either the NEBNext ChIP-seq Library Prep Master Mix Set for Illumina kit (New England Biolabs) or the SPARK DNA Sample Prep Kit Illumina Platform (Enzymatics) according to the manufacturer's instructions.
aggregate value, smoothing of 6, and missing data were considered as “0” (with the exception of Fig. 1a where windows size was set to 50 pb). We used four reference points for most of the figures to avoid contamination from adjacent genes; in such a case there are five blocks of data corresponding respectively to the upstream gene (signal ignore), the upstream intergenic region (signal aligned toward the TSS of the gene of interest), the gene of interest (signal split such that the first half was aligned toward the TSS and the second half toward the 3′end/polyA), the downstream intergenic region (signal aligned toward the 3′end/polyA), and the downstream gene (signal ignore). For figures showing either the average profile of data sets over the complete 4755 mRNA genes with associated PAS (Figs. 1a and 4a), the 31 monocistronic snoRNA genes (Fig. 1b) or the 24 monocistronic snoRNA genes with identified PAS (Fig. 4b), only signal in the first and last blocks were ignored. Similarly, for figures focusing only on the 3′ region of the 4755 mRNA genes with associated PAS (Figs. 1f and 4e), the 31 monocistronic snoRNA genes (Fig. 1g) or the 24 monocistronic snoRNA genes (Fig. 1b) or the 24 monocistronic snoRNA genes (Fig. 4f), only the signal in the second half of the third block (corresponding to the gene of interest) and the fourth block (downstream intergenic region) was shown. To better represent the extension of signal in the mutant strains (Figs. S4–g, 6d, e, and Supplementary Fig. 5a-l-m), we used only two reference points where the first block contains the signal in the upstream intergenic region potentially contaminated with signal from the upstream gene (signal aligned as described previously), the second block contains the signal of the gene of interest, and the third block the downstream intergenic region also potentially contaminated with signal from the downstream gene. Genome-wide Pearson correlation coefficients (Fig. 4g and Supplementary Fig. 4) were calculated using the epiGeEC tool version 1.02.

Code availability
All scripts used for data processing and statistical analysis were written in Python, Perl, or R, and are available upon request.

Data availability
Sequencing data sets have been deposited in the GEO database under accession numbers GSE115595 and GSE95139 for ChIP-seq and 3′ end RACE, respectively. All other data generated or analyzed during this study are included in this published article and its Supplementary Information or from the authors on reasonable request.

Received: 23 February 2018 Accepted: 30 August 2018
Published online: 19 October 2018

References
1. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. Science 352, aa9926 (2016).
2. Porrua, O. & Libri, D. Transcription termination and the control of the transcriptome: why, where and how to stop. Nat. Rev. Mol. Cell Biol. 16, 190–202 (2015).
3. Casanal, A. et al. Architecture of eukaryotic mRNA 3′-end processing machinery. Science. https://doi.org/10.1126/science.aao6535 (2017).
4. Shi, Y. & Manley, J. L. The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev. 29, 889–897 (2015).
5. Baejen, C. et al. Architecture of eukaryotic mRNA 3′-end processing machinery. Science. https://doi.org/10.1126/science.aao6535 (2017).
6. Baejen, C. et al. Genome-wide analysis of RNA polymerase II termination by binding RNA polymerase II and nascent RNA. Nat. Commun. 8, 14861 (2017).
7. Banerjee, A., Apponi, L. H., Pavlath, G. K. & Corbett, A. H. PABPN1: molecular function and muscle disease. FEBS J. 280, 4230–4250 (2013).
8. Larochelle, M., Lemay, J. F. & Bachand, F. The THO complex cooperates with the nuclear RNA surveillance machinery to control small nuclear RNA expression. Nucleic Acids Res. 40, 10240–10253 (2012).
9. Larochelle, M. et al. The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nuclear RNAs. Mol. Cell 37, 34–45 (2010).
10. Lemay, J. F. et al. The Nrd1-like protein Srb1 coordinates cotranscriptional 3′ end processing and polyadenylation site selection. Genes Dev. 30, 1558–1572 (2016).
11. Wittmann, S. et al. The conserved protein Srb1 drives transcription termination by binding RNA polymerase II and nascent RNA. Nat. Commun. 8, 14861 (2017).
12. Carroll, K. L., Pradhan, D. A., Granek, J. A., Clarke, N. D. & Corden, J. L. Polyadenylation of yeast 3′-end RNA processing endonuclease. FEBS J. 280, 4230–4250 (2013).
13. Larochelle, M., Hunyadkurti, J. & Bachand, F. Polyadenylation site selection: linking transcription and RNA processing via a conserved carboxy-terminal domain (CTD)-interacting protein. Curr. Genet. 63, 195–199 (2017).
14. Wang, L., Dao, D., Weisman, R. & Forsburg, S. L. Rapid regulation of nuclear proteins by rapamycin-induced translocation in fission yeast. Yeast 31, 253–264 (2014).
15. Hoque, M. et al. Analysis of alternative cleavage and polyadenylation by 3′-end region extraction and deep sequencing. Nat. Methods 10, 133–139 (2013).
16. Liu, X. et al. Comparative analysis of alternative polyadenylation in S. cerevisiae and S. pombe. Genome Res. 27, 1685–1695 (2017).
17. Rimal, N. et al. Genome-wide mapping of polyadenylation sites in fission yeast reveals widespread alternative polyadenylation. RNA Biol. 10, 1407–1414 (2013).
18. Schrack, M. et al. Genome-wide analysis of poly(A) site selection in Schizosaccharomyces pombe. RNA 19, 1617–1631 (2013).
19. Jimeno-Gonzalez, S., Schmid, M., Malagon, F., Haaning, L. J. & Jensen, T. H. Rat1p maintains RNA polymerase II CTD phosphorylation balance. RNA 20, 531–558 (2014).
20. Orlando, D. A. et al. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. Cell Rep. 9, 1163–1170 (2014).
21. Mandel, C. R. et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end processing endonuclease. Nature 444, 953–956 (2006).
22. Jeronimo, C. & Robert, F. Kim28 regulates the transient association of mediator with core promoters. Nat. Struct. Mol. Biol. 21, 449–455 (2014).
23. Ryan, K., Calvo, O. & Manley, J. L. Evidence that polyadenylation factor CPSF-73 is the mRNA 3′ processing endonuclease. RNA 10, 565–573 (2004).
24. Marguerat, S. et al. Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. Cell 151, 671–683 (2012).
25. Touat-Todeschini, L. et al. Selective termination of Inr RNA transcription promotes heterochromatin silencing and cell differentiation. EMBO J. 36, 2655–2664 (2017).
transcription and the localization of condensin at Pol III genes. PLoS Genet. 10, e1004794 (2014).
45. Shoucri-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell. 42, 794–805 (2011).
46. Zhao, D. Y. et al. SMN and symmetric arginine dimethylation of RNA polymerase II CTD regulate domain control termination. Nature 529, 48–53 (2016).
47. Schaughency, P., Merran, J. & Corden, J. L. Genome-wide mapping of yeast RNA polymerase II termination. PLoS Genet. 10, e1006632 (2014).
48. Grzechnik, P., Gdula, M. R. & Proudfoot, N. J. Pcf11 orchestrates transcription termination pathways in yeast. Genes Dev. 29, 849–861 (2015).
49. Kim, M. et al. Distinct pathways for snoRNA and mRNA termination. Mol. Cell 74, 723–736 (2019).
50. Nedea, E. et al. The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. Mol. Cell 29, 577–587 (2008).
51. Prade-Balade, B. et al. CRM1 controls the composition of nucleoplasmic pre-snRNA complexes to licence them for nucleolar transport. EMBO J. 30, 2212–2221 (2011).
52. Lemieux, C. & Bachand, F. Cotranscriptional recruitment of the nuclear poly(A)-binding protein Pab2 to nascent transcripts and association with translating mRNPs. Nucleic Acids Res. 37, 3418–3430 (2009).
53. Shi, Y. et al. Molecular architecture of the human pre-mRNA 3′ processing complex. Mol. Cell 33, 365–376 (2009).
54. Jamonnik, N. et al. Yeast Nrd1, Nab3, and Sen1 transcription-replicating maps suggest multiple roles in post-transcriptional RNA processing. RNA 17, 2011–2025 (2011).
55. Nabavi, S. & Nazar, R. N. Cleaveage-induced termination in U2 snRNA gene expression. Biochem. Biophys. Res. Commun. 393, 461–465 (2010).
56. O’Reilly, D. et al. Human snRNA genes use polyadenylation factors to promote efficient transcription termination. Nucleic Acids Res. 42, 264–275 (2014).
57. Cassart, C., Drogat, J., Migeot, V. & Hermand, D. Distinct requirement of RNA polymerase II CTD phosphorylations in budding and fission yeast. Transcription 3, 231–234 (2012).
58. Schwer, B. & Shuman, S. Deciphering the RNA polymerase II CTD code in fission yeast. Mol. Cell 43, 311–318 (2011).
59. Battaglia, S. et al. RNA-dependent chromatin association of transcription elongation factors and Pol II CTD kinases. elife 6, e25637 (2017).
60. West, M. L. & Corden, J. L. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. Genetics 140, 1223–1233 (1995).
61. Jimeno-Gonzalez, S., Haaning, L. L., Malagon, F. & Jensen, T. H. The yeast 5′-3′ exonuclease Rat1p functions during transcription elongation by RNA polymerase II. Mol. Cell 37, 580–587 (2010).
62. Nguyen, D. et al. Polyadenylation-dependent 3′ end maturation pathway is required for the synthesis of the human telomerase RNA. Cell Rep. 13, 2244–2257 (2015).
63. Bahlar, J. et al. Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast 14, 943–951 (1998).
64. Waddle, J. A., Karpova, T. S., Waterston, R. H. & Cooper, J. A. Movement of cortical actin patches in yeast. J. Cell Biol. 132, 861–870 (1996).
65. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).
66. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595 (2010).
67. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).
68. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics 26, 2204–2207 (2010).
69. Casper, J. et al. The UCSC genome browser database: 2018 update. Nucleic Acids Res. 46, D762–D769 (2018).
70. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Meth. 9, 357–359 (2012).
71. McDowell, M. D. et al. Pombase 2015: updates to the fission yeast database. Nucleic Acids Res. 43, D656–D661 (2015).
72. Wood, V. et al. PomBase: a comprehensive online resource for fission yeast. Nucleic Acids Res. 40, D695–D699 (2012).
73. Brunelle, M. et al. Aggregate and heatmap representations of genome-wide localization data using VAP, a versatile aggregate profiler. Methods Mol. Biol. 1334, 273–298 (2015).
74. Coulombe, C. et al. VAP: a versatile aggregate profiler for efficient genome-wide data representation and discovery. Nucleic Acids Res. 42, W485–W493 (2014).
75. Laperle, J. et al. The epigenomic efficient criterion (epGeC) tool allows fast comparison of user datasets with thousands of public epigenomic datasets. Bioinformatics, https://doi.org/10.1093/bioinformatics/bty655 (2018).

Acknowledgements
We thank Michelle Scott and François Robert for critical reading of the manuscript; Luc Gaudreau for S. cerevisiae strains; the sequencing platforms of the McGill University and Genome Québec Innovation Centre; Calcul Québec and Compute Canada provided the computing infrastructure used to analyze the data. This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) to F.B. (RGPIN-2017–05482) and P.-É. J. (435710–2013), and by funding from the National Institute of General Medical Sciences to B.T. (GM084089); P.-E.J. is supported by the Fonds de recherche du Québec – Santé (FRQS) and F.B. holds a Canada Research Chair in Quality Control of Gene Expression.

Author contributions
M.L. and F.B. conceived the study and the experimental frame, while M.-A.R. and P.-E.J. planned and designed the pipelines for processing the genome-wide data. M.I. prepared chromatin extracts for ChIPs, performed ChIP-seq, including QCs and library preparations. M.L. performed most of the RNA analyses with help from J.-N.H., J.-N.H. made the Ysh1 constructs and performed the phenotypic characterization of the ysh1 and ma14 anchor-away mutants. X.L. prepared the 3′READS libraries with help from B.T., D.M. and S.R. helped with the preparation of the CHIP-seq libraries. M.-A.R. and P.-E.J. performed all of the bioinformatics analyses of the CHIP-seq and 3′READS data with the help of X.L., M.L., M.-A.R., P.-E.J. and F.B. prepared and finalized the figures. F.B. wrote the manuscript with help of P.-E.J., which was reviewed by all authors.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06546-x.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.