Edinburgh Research Explorer

Surveillance-ready transcription

Author for correspondence:
David Tollervey
e-mail: d.tollervey@ed.ac.uk

Surveillance-ready transcription: nuclear RNA decay as a default fate

Stefan Bresson and David Tollervey

Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh, EH9 3BF, UK

DT, 0000-0003-2894-2772

Eukaryotic cells synthesize enormous quantities of RNA from diverse classes, most of which are subject to extensive processing. These processes are inherently error-prone, and cells have evolved robust quality control mechanisms to selectively remove aberrant transcripts. These surveillance pathways monitor all aspects of nuclear RNA biogenesis, and in addition remove nonfunctional transcripts arising from spurious transcription and a host of non-protein-coding RNAs (ncRNAs). Surprisingly, this is largely accomplished with only a handful of RNA decay enzymes. It has, therefore, been unclear how these factors efficiently distinguish between functional RNAs and huge numbers of diverse transcripts that must be degraded. Here we describe how bona fide transcripts are specifically protected, particularly by 5′ and 3′ modifications. Conversely, a plethora of factors associated with the nascent transcripts all act to recruit the RNA quality control, surveillance and degradation machinery. We conclude that initiating RNAPII is 'surveillance ready', with degradation being a default fate for all transcripts that lack specific protective features. We further postulate that this promiscuity is a key feature that allowed the proliferation of vast numbers of ncRNAs in eukaryotes, including humans.

1. Introduction

Almost all RNA species undergo elaborate maturation processes within the nucleus. In the case of messenger RNAs (mRNAs), nascent transcripts are synthesized by RNA polymerase II (RNAPII) as pre-mRNAs consisting of both introns and exons (figure 1). Introns are removed and generally degraded, while the intervening exons are spliced together to generate the mature message. Further modifications are made to the ends of the transcript. An inverted, 7-methylguanosine cap structure is added to the 5′ end, and a poly-adenylated (poly(A)) tail is synthesized at the 3′ end. In parallel, RNA binding proteins package the transcripts into export-competent mRNP particles, which are subsequently transported to the cytoplasm. These events are coordinated by the repetitive carboxy-terminal domain (CTD) of RNAPII, which acts as a general binding platform for RNA processing factors (reviewed in [1]). The heptad repeats (26 in yeast and 52 in humans) of the CTD are differentially phosphorylated throughout the transcription cycle, allowing distinct sets of maturation factors to be recruited at the correct time and place.

The complexity of nuclear RNA processing makes it inevitable that some fraction of nascent transcripts will fail to mature correctly. The accumulation of aberrant or defective transcripts represents a significant potential problem, because they could saturate the RNA processing machinery and impede the production of functional products. For example, accumulation of cryptic RNAs in yeast mutants with defective RNA degradation reduces the availability of the nuclear cap binding complex, with pleiotropic effects on gene expression [2]. Antisense transcripts can hybridize to complementary sense RNA, forming double-stranded RNAs that may enter the RNA interference pathway [3]. Excess RNA, particularly if poorly packaged, can also bind to homologous
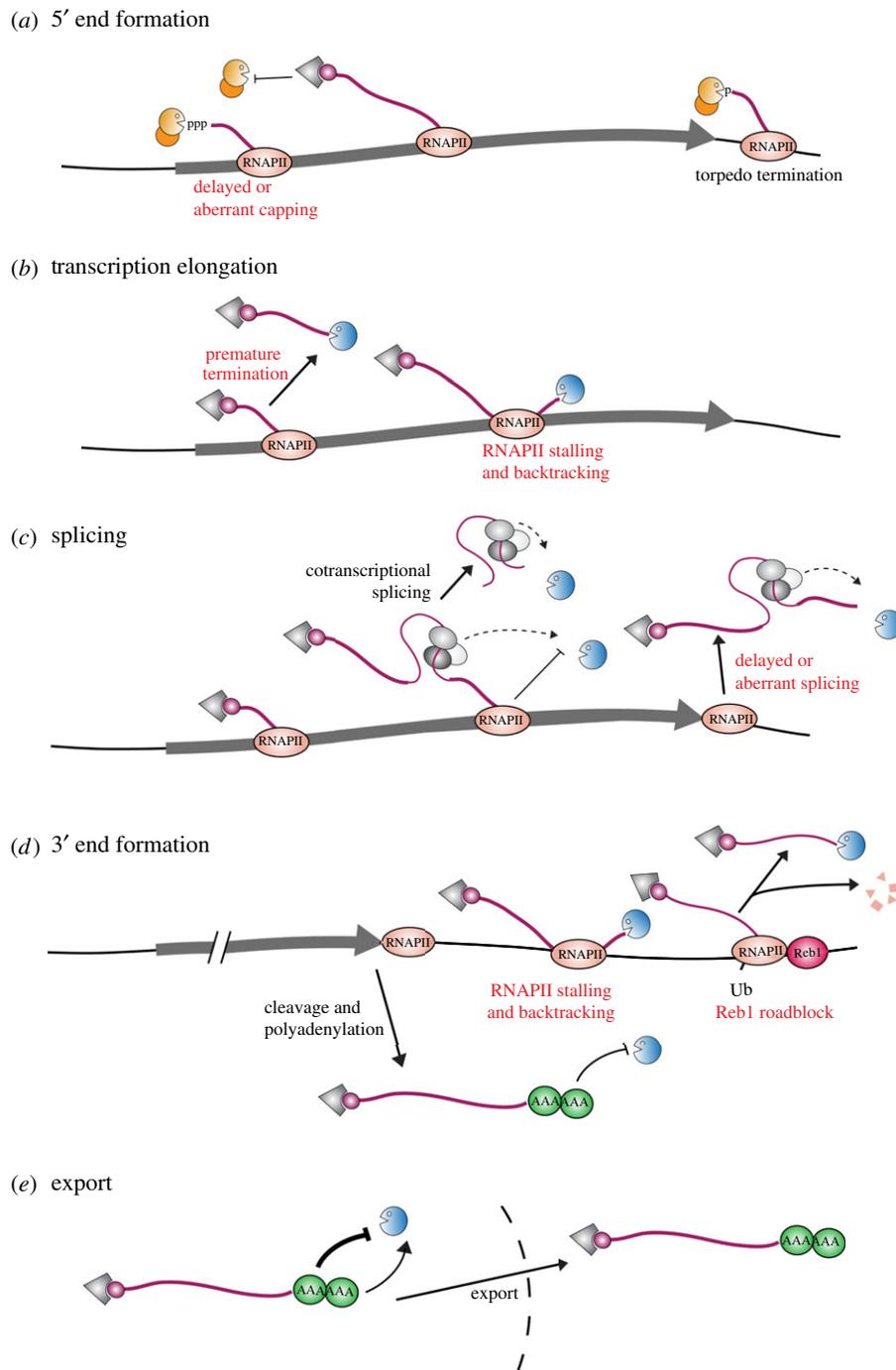
Figure 1. Processing and surveillance of pre-mRNAs. Multiple steps during mRNA transcription and processing are screened by surveillance activities. (*a*) Delayed or aberrant capping leads to decay by nuclear 5′ surveillance pathways. Degradation requires a pyrophosphatase activity (orange circle) to remove the triphosphate and a coupled 5′–3′ exonuclease (orange pacman). Correctly maturing transcripts are protected by the presence of the m^{7}G cap and the cap binding complex (CBC; grey triangle). Following normal transcript cleavage and polyadenylation, the 3′ fragment of the nascent transcript is targeted by the 5′ exonuclease in order to terminate RNAPII transcription. (*b*) Prematurely terminated transcripts are 3′ degraded by the nuclear exosome (blue pacman). Transcription termination and surveillance can involve either complete dissociation of the polymerase (*left*) or polymerase backtracking to reveal the 3′ end, providing an entry point for the exosome (*right*). (*c*) Unspliced transcripts are targeted by the surveillance machinery. In normal mRNA biogenesis, introns are typically spliced cotranscriptionally. Excised introns must be constitutively degraded and features associated with splicing or introns may act to recruit the nuclear surveillance machinery. When introns are not efficiently removed, these factors may facilitate degradation of the entire transcript. (*d*) Aberrant 3′ end formation leads to surveillance by the nuclear exosome. In fission yeast, this can involve RNAPII stalling and backtracking downstream of the PAS (centre). Alternatively, the budding yeast protein Reb1 (red circle) can terminate transcription by functioning as a roadblock (*right*). RNAPII is ubiquitinated and degraded, and the released transcript is degraded by the nuclear exosome. Correctly terminated transcripts (*left*) are protected by a poly(A) tail appropriately packaged with poly(A) binding proteins (green circles). (*e*) Transcripts with prolonged nuclear retention are subject to slow, default surveillance pathways. This process appears to be facilitated in part by nuclear poly(A) binding proteins, which protect the transcript but can also stimulate decay through recruitment of the nuclear surveillance machinery.

gene loci, forming harmful RNA:DNA hybrids that are associated with DNA double-strand breaks. In the cytoplasm, aberrant mRNAs may encode truncated, nonfunctional or even dominant negative proteins.

Historically, RNA quality control or 'surveillance' mechanisms have been difficult to examine in unperturbed systems, as defective transcripts generally constitute only a small proportion of any given endogenous mRNA species.
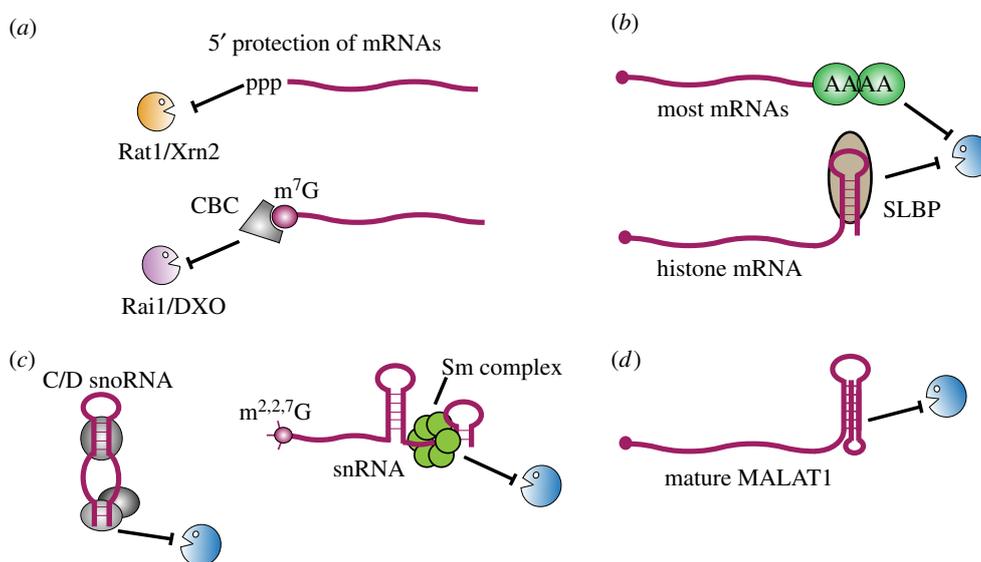
Figure 2. Protective features in RNA stability. (*a*) Primary RNAPII transcripts are initially protected by the terminal 5′ triphosphate, which blocks degradation by the nuclear 5′−3′ exonucleases Rat1/Xrn2. Transcripts are generally rapidly modified by addition of an inverted GpppN cap structure. This is sensitive to removal by the pyrophosphatases Rai1 and DXO, but undergoes m7G methylation and association with the cap binding complex (CBC), conferring pyrophosphatase resistance. (*b*) Most mRNAs are shielded at their 3′ end by a poly(A) tail packaged with poly(A) binding proteins. The non-polyadenylated, replication-dependent histone mRNAs are protected by a terminal stem−loop structure bound to the stem−loop binding protein (SLBP). (*c*) Small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs) are shielded from exosome-mediated decay by specific proteins bound to the 3′ end. snRNAs and many snoRNAs are protected at their 5′ end by the trimethylated $m^{2,2,7}G$ cap. (*d*) The mature MALAT1 transcript contains a triple helix that sequesters the 3′ end and prevents 3′−5′ exonucleolytic decay.

Most studies have, therefore, relied on mutations, either in reporter transcripts or the RNA processing machinery, to artificially trigger surveillance. These analyses suggest that the surveillance machinery monitors a wide range of processing defects, including transcripts with defects in cap structure [4]; inefficient polyadenylation [5]; aberrant splicing or 3′ end formation; improper mRNP packaging [6]; or inefficient nuclear export (figure 1). These defects have little in common, but all apparently lead to recognition and destruction of the RNA by the surveillance machinery. Moreover, the same core system degrades transcripts generated by RNA polymerases I and III, which are significantly different in structure and packaging from most RNAPII products.

Surveillance pathways also degrade RNAs that result from pervasive transcription (reviewed in [7]). Eukaryotic promoters generally drive transcription initiation in both directions, but in most cases only one side results in productive gene expression [8–11]. In part, directionality is enforced through selective degradation of the upstream antisense transcript [12–16]. These RNAs are referred to as cryptic unstable transcripts (CUTs) in yeast, and promoter upstream transcripts (PROMPTs) or upstream antisense RNAs (uaRNAs) in mammalian cells. While bidirectional promoters are a prominent source of transcriptional noise, many or all active enhancer elements are also transcribed and cryptic transcription can initiate from any nucleosome free region [9]. The resulting transcripts are highly unstable in wild-type cells but accumulate when the surveillance machinery is inactive [13,16,17].

In organisms as diverse as yeast, plants and humans, RNA surveillance depends on the nuclear exosome, a complex with endonuclease and 3′ exonuclease activity. The exosome, in turn, relies on numerous cofactors to guide it to target transcripts and help initiate decay. Interestingly, many of these cofactors also function in RNA maturation, suggesting that the exosome is recruited to nascent transcripts regardless of

their processing status. As discussed below, the exosome appears to act as a general scavenger of 3′ ends, potentially degrading nuclear pre-mRNAs and other transcripts by default. Correctly processed transcripts largely escape nuclear surveillance through the deposition of specific RNA binding proteins, particularly at the 3′ end, which sterically hinder exonucleolytic decay.

In this review, we discuss quality control mechanisms for RNAPII transcripts in *Saccharomyces cerevisiae* and human cells, with additional reference to RNAPI and studies in other species as appropriate. We begin with an introduction to the major cellular exonucleases and cofactors, highlighting their parallel roles in mRNA biogenesis and surveillance. Subsequently, we review mechanisms by which the exosome degrades aberrant transcripts, with a particular focus on the connections between transcription termination and surveillance.

2. Surveillance machinery

A striking feature of eukaryotic RNA degradation and surveillance pathways is the preponderance of exonucleases, which degrade RNAs from the 5′ end (5′ exonucleases) or 3′ end (3′ exonucleases), rather than endonucleases that can cleave RNAs internally. As a consequence, accessibility of the 5′ or 3′ ends of the transcript for nuclease attack is likely to be a key feature in determining susceptibility to degradation (figure 2).

2.1. 5′ Exonucleases

The major 5′ exonucleases in eukaryotes are the related proteins Xrn1, which is predominantly cytoplasmic, and Rat1 (Xrn2 in humans), which is predominantly nuclear. The activities of both enzymes are largely blocked by the presence of a 5′ triphosphate [18], which is initially present on all newly synthesized transcripts. This protection can be

seen during degradation of the excised 5′ external transcribed sequence (5′ETS) spacer of the pre-ribosomal RNA (pre-rRNA). The 5′ region, which carries the 5′ triphosphate, is degraded by the exosome (see below), whereas the 3′ region, which is generated by cleavage and carries a 5′ mono-phosphate, is degraded by Rat1 [19,20]. Similarly, 3′ cleavage of the pre-rRNA allows entry of Rat1, which degrades the downstream nascent transcripts. The pre-mRNA 3′ cleavage and polyadenylation machinery also leaves a 5′ monophosphate, allowing Rat1/Xrn2 to degrade the downstream nascent transcript [19–21]. Transcription is then terminated by Rat1/Xrn2, presumably acting when it catches the transcribing polymerase, in a phenomenon referred to as 'torpedo' termination [19–21].

Nascent RNAPII transcripts are further protected by addition of a 7-methylguanosine (m^7G) cap to the 5′ triphosphate end of the RNA. This reaction occurs shortly after transcription initiation, usually within the first 50 nt. Capping defects can be induced by inactivation of the capping machinery itself, or indirectly through mutations in RNAPII [22]. Rat1 forms a complex with the pyrophosphatase Rai1 and transcripts that fail to be capped, or on which the cap is not 7-methylated, are rapidly decapped and dephosphorylated by Rai1 [23]. This exposes the transcript to degradation by Rat1 [19,22,24–26]. Capped transcripts are further protected by the nuclear cap binding complex (CBC), comprising Cbc1–Cbc2 in yeast and CBP80–CBP20 in humans. CBC directly blocks access to the 5′ cap by the decapping enzyme Dxo1, which is homologous to Rai1 and can also initiate degradation [27,28].

Since decapping leaves a 5′ monophosphate, it seems likely that RNAPII will also be subject to torpedo termination by Rat1 when capping is defective. Notably, the sensitivity of degradation systems to 5′ nucleotide status is conserved in evolution, because the major bacterial endonuclease RNase E is active only on substrates with a 5′ monophosphate [29].

2.2. The exosome

In eukaryotes, the major 3′ exonuclease activity during RNA surveillance is supplied by the nuclear exosome complex. The exosome was originally defined for its role in the processing of precursors to stable RNAs; rRNA, small nucleolar RNA (snoRNA) and small nuclear RNA (snRNA) are all processed at least in part by the exosome [30,31]. Subsequently, the exosome was shown to target a wide variety of transcripts, including defective pre-rRNAs, pre-tRNAs, aberrant mRNAs and transcripts arising from pervasive transcription [13,16,17,32–36]. Exosome structure and function is the subject of several excellent reviews [37–39], and will be discussed only briefly here.

The exosome consists of a central hexameric ring comprising six proteins, with an additional three subunits layered on top. These nine proteins are arranged in a barrel surrounding a central channel just wide enough to accommodate single-stranded RNA. Collectively, this complex is referred to as the exosome core, and is structurally conserved in Archaea. In eukaryotes, catalytic activity is supplied by the associated, highly processive 3′ to 5′ exonuclease Dis3/Rrp44. Dis3 is positioned at the base of the barrel, and approximately 30 nt of single-stranded RNA must be threaded through the central channel to reach the active site [40,41]. Dis3 has an additional PIN endonuclease domain, but the range of

in vivo targets for this endonuclease activity is unclear; the only confirmed targets are the 5′ETS and pre-5.8S regions of the pre-rRNA. The nuclear exosome can also associate with the distributive 3′ to 5′ exonuclease Rrp6, assisted by its cofactor Rrp47 (C1D in human cells) [42–44]. Dis3 and Rrp6 show some functional specialization, but there is considerable overlap in their target sets [32,33].

2.3. Nuclear exosome cofactors; TRAMP, NNS, NEXT

The eukaryotic exosome core is relatively inactive, making its function heavily reliant on cofactors. A well-studied cofactor is the Trf4/5–Air1/2–Mtr4 polyadenylation (TRAMP) complex. In *S. cerevisiae*, TRAMP consists of a poly(A) polymerase (PAP) (either Trf4 or Trf5), a zinc-finger, RNA binding protein (Air1 or Air2) and the RNA helicase Mtr4 [16,45,46]. Trf4/5 adds a short oligo(A) tail to the transcript end, which Mtr4 threads into the exosome central channel. Air1 and Air2 are thought to aid in substrate binding. *In vivo* cross-linking experiments with Dis3 reveal a high fraction of reads with non-templated A-tails, suggesting that TRAMP-mediated polyadenylation generally precedes exosome-mediated decay [33]. Indeed, TRAMP plays an essential role in the degradation of nearly all surveillance targets of the exosome in yeast, including defective, hypo-modified pre-tRNA$_i^{Met}$ [47], defective pre-rRNAs [48] and cryptic RNAPII transcripts [16].

The role of Mtr4 and the exosome in ribosome production is complex. Mtr4 acts in 'constitutive' pre-rRNA processing steps: 3′ trimming of the pre-5.8S rRNA and turnover of the excised 5′ETS spacer. Specific adaptor proteins, the ribosome biogenesis factors Nop53 and Utp18, interact with the arch domain of Mtr4 during recruitment to the pre-5.8S rRNA and 5′ETS, respectively [49]. Notably, both the arch domain of Mtr4 and the arch-interacting motif of Nop53 are con-served in higher eukaryotes. When ribosome synthesis is proceeding normally, exosome activity is terminated at the 5′ end of the pre-rRNA or close to the mature 3′ end of the 5.8S rRNA. However, on pre-rRNAs that fail to undergo cotranscriptional cleavage, recruitment of the exosome leads to degradation of the entire pre-rRNA.

There is little evidence that the pre-rRNA processing steps involve the TRAMP complexes, which probably function only in pre-rRNA surveillance. Consistent with this, loss of Trf5 was shown to partially rescue ribosome synthesis in several different pre-ribosome assembly mutants [50,51]. Moreover, hyperadenylated pre-rRNAs accumulate on depletion of Mtr4 [52], showing that the Trf–Air module of TRAMP can be recruited independently of Mtr4 and, presumably, of the exosome [51,52].

Notably, the processing roles of Mtr4 and exosome take place in the context of a specific, on-pathway pre-ribosomal particle, in which the interactions have presumably been fine-tuned through evolution. During surveillance, in contrast, TRAMP, Mtr4 and the exosome must respond to a multitude of different, off-pathway particles with diverse defects in RNA processing, RNA folding and RNA–protein interactions. Similar issues must arise during surveillance of other RNA classes; the list of defects that could potentially arise seems almost endless. We therefore predict that the surveillance system is keyed to recog-nize 'generic' attributes, and uses general adaptor proteins, rather than recruitment by specific features of any particular, misassembled RNA–protein complex.

TRAMP-mediated surveillance of RNAPII transcripts is usually coupled to transcription termination by the Nrd1–Nab3–Sen1 (NNS) complex [14,15,53]. Nrd1 and Nab3 are RNA binding proteins that each bind a short consensus motif [54,55]. However, these sequence elements are abundant throughout the genome and usually insufficient to drive binding [56]. Nrd1–Nab3 recruitment is aided by a physical interaction between Nrd1 and phosphorylated serine 5 (Ser5-P) within the CTD of RNAPII [57,58] (figure 3). Ser5-P is ubiquitously present during early transcription and, probably as a result, Nrd1 and Nab3 are broadly recruited to the 5′ ends of nascent transcripts [54–56,59]. Once bound, Nrd1 and Nab3 can engage Sen1, an ATP-dependent, 5′–3′ RNA and DNA helicase. Sen1 is the putative transcription termination factor, which may displace polymerase and free the nascent transcript for degradation by the TRAMP–exosome complex [14,15,53,60–62]. Termination by NNS is further promoted by interactions with histone H3 modified by lysine 4 trimethylation (H3K4me3) [63], a hallmark of the 5′ regions of protein-coding genes.

The NNS pathway also participates in the maturation of snoRNAs [64], promoting 3′ processing rather than complete degradation. This is possible because specific snoRNA-binding proteins mark the mature 3′ end and block further degradation by the exosome.

The NNS pathway seems to be unique to budding yeasts; there are apparent human homologues of Nrd1 (SCAF8/RBM16) and Sen1 (Senataxin), but these are not known to form a complex with a Nab3 homologue. By contrast, the TRAMP complex is well conserved [65], but human TRAMP is confined to the nucleolus, where it presumably participates in pre-rRNA surveillance [66]. hMtr4 is distributed more broadly throughout the nucleus, and can also associate with the RNA binding proteins RBM7 and ZCCHC8 to form the nuclear exosome targeting (NEXT) complex [66]. Both RBM7 and ZCCHC8 are restricted to the nucleoplasm, and NEXT participates in surveillance of RNAPII transcripts, including the degradation of cryptic transcripts. Additionally, NEXT facilitates the termination and 3′ end processing of human snRNAs [66,67]. NEXT is recruited to target transcripts through interactions with the nuclear CBC [68] and, therefore, binds promiscuously during early RNAPII transcription (figure 3). Subsequently, the CBC–NEXT complex triggers early termination and degradation of target transcripts [69]. These features are analogous to the NNS pathway, despite the lack of sequence homology between the individual components.

3. Surveillance throughout the transcription cycle

3.1. Transcription elongation

During transcription elongation, the 3′ end of the nascent transcript is buried within the active site of RNAPII and inaccessible to 3′ exonucleases. It seems likely that exosome-mediated surveillance will first require transcription termination and dissociation of the polymerase. Indeed, this appears to be the mechanistic basis for the NNS and NEXT pathways. There is, however, another potential pathway that might allow the exosome access to the 3′ ends of nascent transcripts. When transcription elongation is impeded, the
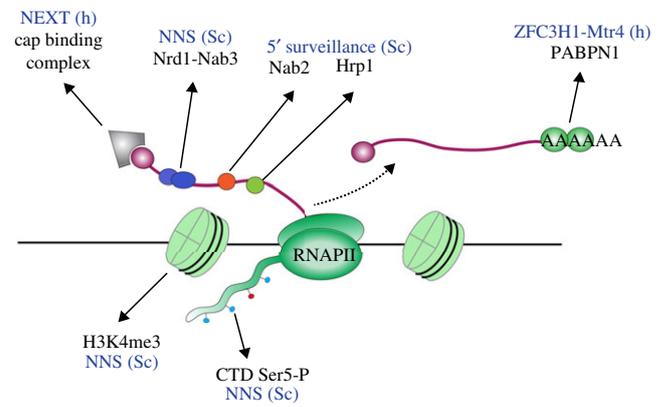

Figure 3. Pro-surveillance factors in RNAPII transcription units. Nascent transcripts contain numerous features which facilitate the recruitment of surveillance factors. Importantly, these factors are also present during normal RNA biogenesis, which presumably allows the nuclear surveillance machinery to 'inspect' all nascent RNAs. Details are discussed in the text. h, human; Sc, *Saccharomyces cerevisiae*.

polymerase can backtrack—sliding backwards along the DNA and extruding the 3′ end of the nascent transcript from the catalytic site (reviewed in [70]). Backtracked RNAPII is generally rescued by transcription factor IIS (TFIIS; Dst1 in yeast) [71], which activates the intrinsic hydrolytic activity of RNAPII. This cleaves the nascent transcript and realigns the 3′ end within the active site, allowing transcription to resume.

In principle, extended backtracking might also lead to the extruded 3′ end projecting from the polymerase to a sufficient extent to allow binding of the surveillance machinery. This could provide an entry point for the nuclear exosome to degrade the nascent transcript, displacing RNAPII in the process [72]. Consistent with this model, depletion of Dis3 in the fission yeast *Schizosaccharomyces pombe* results in extended transcripts with elevated RNAPII occupancy downstream of the poly(A) site [72]. RNAPII is known to pause downstream of the poly(A) site, allowing time for cotranscriptional cleavage and polyadenylation [73,74]. This pause is proposed to favour backtracking, leading to transcription termination via the exosome [72]. For transcripts that are rapidly cleaved and polyadenylated prior to transcription termination, the exosome would degrade only the sequence downstream of the poly(A) site. The remainder of the transcript would be protected due to its separation from the elongation complex by 3′ cleavage. Were RNA cleavage delayed, the exosome could destroy the entire transcript, providing quality control for timely 3′ end formation. RNAs lacking bona fide polyadenylation signals or transcripts which fail to correctly assemble the cleavage and polyadenylation complex would then be degraded by default (figure 1). This model might also rationalize observations from *S. cerevisiae* which implicate the exosome in mRNA transcription termination [32,75]. Backtracking has also been invoked to explain transcription termination in human cells [76].

3.2. Splicing

Unspliced pre-mRNAs are very rapidly degraded when splicing is delayed, indicating active surveillance. However, the mechanism and factors involved remain obscure. In yeast and humans, splicing generally occurs cotranscriptionally,

so introns and splicing factors are usually removed before the 3′ end is exposed by transcription termination. Since excised introns must always be degraded, or trimmed to release internal snoRNAs, it is possible that features or factors associated with splicing would act to specifically recruit the RNA degradation system. When splicing is delayed, the 3′ end is no longer protected by the transcription elongation complex and these factors could facilitate degradation of the entire transcript. Degradation could be induced through direct recruitment of the exosome by one or more splicing factors, or through specific decay-promoting sequences within introns. The clearest example is the binding of Mmi1 to fission yeast introns [77]. Mmi1 recruits the exosome to a subset of unspliced introns, facilitating degradation of the entire RNA. However, these transcripts will generally be terminated by the canonical cleavage machinery, so degradation either outcompetes polyadenylation or is able to overcome the protection a poly(A) tail would normally confer. Whether additional intron-associated factors participate in surveillance is unclear, so this model remains somewhat speculative. However, if multiple exosome recruitment pathways exist, loss of a single factor might not have dramatic effects.

3.3. Termination

Eukaryotic pre-mRNAs are generally terminated following cotranscriptional endonuclease cleavage downstream of the open reading frame. In nearly all instances, pre-mRNA cleavage is coupled to the addition of a poly(A) tail. In contrast to alternative termination mechanisms, pre-mRNA cleavage and polyadenylation is not clearly associated with exosome-mediated processing or decay. This is usually attributed to the protective effects of poly(A) binding proteins (poly(A) BPs), which coat the nascent poly(A) tail during its synthesis and are thought to fend off exonuclease attack [78]. This hypothesis has been most directly tested in *S. cerevisiae*, where rapid depletion of the nuclear poly(A) BP Nab2 dramatically destabilizes newly synthesized, polyadenylated mRNAs [79]. This phenotype is alleviated when the exosome is also inactivated, indicating that Nab2 protects poly(A) tailed transcripts from the nuclear exosome (figure 1). Whether this role for nuclear poly(A) BPs is conserved in higher eukaryotes remains uncertain. Human cells encode multiple, potentially redundant, nuclear poly(A) BPs and their roles in mRNA stability remain to be clarified.

In contrast to the protective effects of canonical cleavage and polyadenylation, transcripts terminated by alternative mechanisms are generally highly unstable. The clearest example is NNS-mediated termination, which usually facilitates termination early in the transcription cycle. However, even full-length transcripts are destabilized if termination is not coupled to polyadenylation and poly(A) BP binding. In *S. cerevisiae*, binding sites for the DNA-binding protein Reb1 are enriched in intergenic regions downstream of poly(A) sites [80]. DNA-bound Reb1 acts as an orientation-sensitive roadblock for RNAPII, which stalls and is ultimately ubiquitinated to induce transcription termination [80]. The nascent RNA is released with a 3′ monophosphate, and rapidly degraded by the TRAMP–exosome complex. In alternative pathways, readthrough pre-mRNA transcripts can be terminated by the NNS complex or following transcript cleavage in double-stranded regions by the RNase III homologue Rnt1 [81,82]. In each case, the terminated

transcripts lack a poly(A) tail and are rapidly degraded by the nuclear exosome.

Underlining the importance of 3′ protection, stable but non-polyadenylated RNAPII transcripts all carry specialized 3′ structures and/or specific RNA 3′-end binding proteins (figure 2). The major form of the highly expressed human long noncoding RNA (lncRNA) MALAT1 has a 3′ terminus that is generated by RNase P cleavage at a tRNA-like element [83]. The resulting transcript ends in an A-rich stretch that interacts with two upstream U-rich elements to form a triple-helix structure, which sequesters the 3′ end of the transcript from exonucleases [84,85]. Interestingly, triple-helix formation can also be used to protect a canonical poly(A) tail. For example, the viral noncoding and nuclear-retained RNA PAN carries internal U-rich elements that sequester the poly(A) tail within a triple helix [86] and protect it from exosome-mediated decay [87,88]. Similar structures have been identified in other transcripts, suggesting triple helices could be a widespread mechanism to protect the poly(A) tails of nuclear-retained transcripts [89,90].

In metazoans, the majority of histone proteins are encoded by replication-dependent histone genes (reviewed in [91]). These histone mRNAs are synchronously up-regulated during S phase to support DNA replication. The transcript 3′ end is formed through endonucleolytic cleavage of the pre-mRNA by CPSF73, guided by the U7 snRNA. CPSF73 is also required for canonical cleavage and polyadenylation, but these histone mRNAs are not polyadenylated. Instead, a highly conserved, terminal stem–loop sequence is specifically bound by the stem–loop binding protein (SLBP), which confers protection against 3′ degradation [83].

In yeast, 3′ processing of most snRNAs or snoRNAs is initiated by cotranscriptional cleavage by the endonuclease Rnt1, a homologue of RNase III, which cleaves both sides of a stem–loop structure. The cleaved pre-sn(o)RNA is probably initially protected by 3′ binding of the La protein homologue Lhp1 and/or the Lsm2-8 complex to oligo(U) tracts [92–95], allowing time for assembly of the sn(o)RNP complexes. Loss of the proteins that bind the snRNA or snoRNA 3′ terminus leads to complete degradation of the mature sn(o)RNA region [95,96] (figure 2).

3.4. mRNP packaging

In principle, poorly packaged transcripts which remain bound to chromatin could base pair with the complementary DNA strand. For example, exosome inactivation in human cells leads to the accumulation of enhancer-associated ncRNAs (eRNAs), which generate RNA:DNA hybrids (R-loops) [97]. In both yeast and humans, R-loops are associated with genome instability but are efficiently removed by the RNase H endonuclease [98–100]. RNase H selectively cleaves RNA that is base-paired with DNA, and presumably provides an entry point for 5′ and 3′ exonucleases to degrade the entire RNA molecule. Thus, RNase H may function as an additional layer of RNA surveillance to ensure transcript quality.

4. Dual roles of nuclear poly(A) binding proteins in stability and decay

Despite its role in mRNA stability (see above), Nab2 has also been implicated in RNA surveillance pathways. Nab2 directs
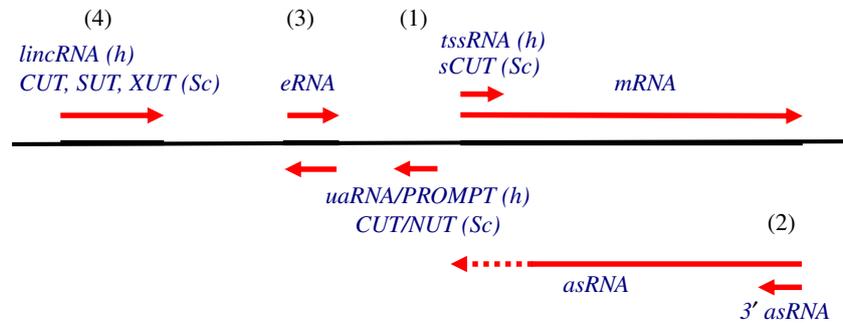
Figure 4. Major classes of noncoding RNA. The regions surrounding eukaryotic protein-coding genes generate a set of ncRNAs in addition to the mRNA transcript. These include: (1) short divergent promoter associated transcripts from the nucleosome-free promoter region, (2) antisense (as) transcripts from the nucleosome-depleted terminator region, (3) enhancer-associated eRNAs. In addition, (4) a range of ncRNAs are transcribed from intergenic locations. h, human; Sc, *Saccharomyces cerevisiae*.

the degradation of pre-mRNAs by Rrp6 [101,102] and autoregulates the *NAB2* mRNA by recruiting the TRAMP–exosome complex [103,104]. Nab2 was also proposed to target CUTs, which generally lack polyadenylation signals and are not polyadenylated [105]. Notably, while Nab2 is enriched at the poly(A) tail, it binds promiscuously throughout transcripts at non-poly(A) sites [105]. These observations indicate that Nab2 can stimulate degradation when not associated with the poly(A) tail.

The human nuclear protein PABPN1 is not related to Nab2 in sequence, but plays an analogous role in nuclear surveillance, targeting a wide range of nuclear substrates. These include spliced genes that encode snoRNAs but no protein product, cryptic transcripts arising from divergent promoters, intronless transcripts and partially spliced mRNAs [88,106–109]. In contrast to Nab2, PABPN1 targets full-length and polyadenylated transcripts, suggesting that PABPN1 stimulates decay while bound to the poly(A) tail. This activity may be enabled by PABPN1's additional role in polyadenylation (reviewed in [78]). On long poly(A) tails, PABPN1 facilitates the distributive extension of the tail by PAP [110], perhaps creating a free 3′ end accessible to the exosome. Consistent with this notion, PAP activity is apparently required for efficient degradation [109].

In addition, PABPN1 has been reported to recruit the zinc-finger protein ZFC3H1, which acts together with Mtr4 to stimulate exosome degradation of the transcript [108,111]. This surveillance pathway is apparently absent from budding yeast but conserved in fission yeast. Red1 is the homologue of ZFC3H1 and works with Mtr4-like 1 (Mtl1) to process snoR-NAs and degrade CUTs, some pre-mRNAs, and meiotic mRNAs, which must be eliminated in mitotic cells [112]. The Mtl1–Red1 core (MTREC) complex is directed to target transcripts by a variety of adaptors, including the PABPN1 homologue Pab2 and the PAP Pla1 [112–117]. The MTREC complex also associates with another 3′ end formation factor Hrp1/Nab4 [117]. In budding yeast Hrp1 is required for pre-mRNA cleavage, binding an upstream UAUAUA sequence [118]. However, like Nab2, Hrp1 also binds at alternative sites and contributes to ncRNA degradation [105] (figure 3).

The dual nature of nuclear poly(A) binding proteins in RNA stability and decay may reflect distinct roles during mRNA biogenesis. Following transcription through the cleavage site, mRNAs are rapidly polyadenylated, a process which is coupled to packaging by poly(A) BPs. Delayed poly(A) binding will expose the 3′ end to exosome-mediated degradation before the transcript can be exported to the cytoplasm. However, because nuclear poly(A) BPs also recruit the

exosome, any transcript with significantly delayed nuclear export will eventually be degraded by default, albeit with much slower kinetics (figure 1). In this respect, it is notable that errors in splicing frequently prevent export, and that Nab2, Pab2 and PABPN1 each target intron-containing mRNAs [101,109,112,113]. In human cells, sensitivity to PABPN1-mediated decay is indeed correlated with prolonged nuclear retention [109,119].

5. Controlling pervasive transcription

Arguably the most significant role of the nuclear surveillance machinery is the removal of huge numbers of different RNAs that arise from pervasive transcription. This term was initially coined in response to the unexpected finding that a large fraction of the human genome is transcribed into unstable RNAs with little or no protein-coding potential [120]. Subsequently, several different classes of ncRNA have been defined and included under this general heading (figure 4). In general, noncoding regions diverge rapidly during evolution, presenting clear problems for their efficient recognition by highly conserved, protein-based surveillance systems. As we discuss below, the surveillance machinery seems to recognize common features of ncRNA biogenesis rather than sequence.

Genome-wide studies indicate that eukaryotic promoters are intrinsically bidirectional [8–11,121]. At many promoters, polymerase initiates transcription equally in both directions, but transcripts in the 'upstream' direction are quickly eliminated by the nuclear surveillance machinery. In both yeast and humans, this surveillance is intimately linked to the mechanism of premature transcription termination.

In *S. cerevisiae*, promoter-associated ncRNA transcription is terminated through the NNS pathway, and the resulting transcripts are degraded by the TRAMP–exosome complex. Nrd1 and Nab3 binding motifs are enriched in transcripts generated upstream of bidirectional promoters, also termed NUTs, and depleted within protein-coding transcripts [53]. This asymmetry presumably favours NNS-dependent termination in upstream regions, but whether it is sufficient to explain the difference in stability between CUTs/NUTs and mRNAs is less clear. Nrd1 and Nab3 binding elements are of low complexity and generally abundant even in protein-coding genes [54–56]. Moreover, cross-linking experiments suggest Nrd1 and Nab3 bind promiscuously to most RNAPII transcripts [53–56].

An additional determinant may be the phosphorylation status of RNAPII. Early transcription is characterized by

high Ser5 and low Tyr1, Ser2 and Ser7 phosphorylation levels along the CTD [122,123]. This combination of marks, termed 'initiation state', helps recruit the NNS complex to nascent transcripts [57,58,122]. For most genes, RNAPII remains in initiation state through the first approximately 150 nt of RNA, roughly corresponding to the first nucleosome [123]. Within this window, polymerase is particularly prone to premature termination and surveillance [105,123]. As transcription continues, the CTD is gradually reconfigured to an elongation state [123]. Most notably, Ser5 phosphorylation is reduced, rendering the transcript less prone to NNS recruitment and transcription termination. Simultaneously, an increase in Ser2 phosphorylation facilitates the recruitment of canonical cleavage and polyadenylation factors [124].

Remarkably, this transition apparently fails to take place during CUT transcription, with polymerase instead remaining stuck in initiation state [123], providing an extended window for NNS-dependent termination. Moreover, the failure to transition to elongation state prevents the recruitment of factors that might otherwise facilitate termination at cryptic poly(A) signals and yield a stable transcript. These factors imply that CUTs are in some sense 'fated' for degradation by the surveillance machinery.

Analogous mechanisms control divergent transcription in humans. Regions upstream of bidirectional promoters are enriched for consensus poly(A) signal (PAS) motifs, aiding early termination by the conventional cleavage and polyadenylation machinery [12,125]. These prematurely terminated transcripts are rapidly degraded by the nuclear exosome, typically in concert with the NEXT complex [66,68]. When present in the sense direction, PAS motifs are usually suppressed by the presence of U1 snRNP along the RNA [125,126]. Together, these features constitute the 'U1–PAS axis' [125], and help enforce transcription directionality. However, even sense transcripts are sometimes subject to premature cleavage and degradation [127].

At first glance, it is surprising that conventional termination should be coupled to decay. However, despite being processed by the canonical cleavage and polyadenylation machinery, uaRNAs are often not polyadenylated [108]. This suggests divergent transcripts are targeted soon after transcript cleavage, perhaps during an early, distributive phase of polyadenylation when the transcript is still unprotected by nuclear poly(A) BPs. This raises the obvious question of why mRNAs are not also susceptible to NEXT-mediated decay. The answer has been proposed to relate to differences in transcript length between uaRNAs and mRNAs [108]. NEXT is recruited by the CBC and, like NNS, preferentially localizes near the 5′ ends of nascent RNAs [68,69]. Cleavage in this region could favour decay triggered by NEXT and thus disfavour polyadenylation. Consistent with this hypothesis, NEXT is most active on transcripts shorter than 2 kb [108]. When transcript cleavage happens far downstream, as is usually the case for mRNAs, polyadenylation should be favoured.

As with yeast, it is unlikely that any one characteristic is sufficient for the mammalian surveillance machinery to distinguish mRNAs from transcriptional noise. Another important feature may be splicing, which has long been associated with RNA stability and surveillance. In mammalian cells, intronless transcripts tend to be weakly expressed and unstable [119] and, while nearly all mammalian mRNAs are spliced, the vast majority of divergent transcripts are not. The connection

between a lack of splicing and uaRNA instability may be direct, because splicing deposits the mRNA export factor REF on nascent transcripts, where it physically interacts with CBC [128]. Notably, REF directly competes with Mtr4 for access to CBC [129], indicating that the export machinery antagonizes nuclear surveillance. Moreover, REF overexpression is sufficient to protect some divergent transcripts from exosome-mediated decay [129]. Somewhat surprisingly, these effects are independent of the NEXT complex, suggesting Mtr4 can also act independently or with other exosome adaptors to degrade divergent transcripts. A plausible candidate is the PABPN1–PAP–ZFC3H1 pathway, which helps to degrade a subset of uaRNAs [107–109].

6. Slow, default decay for nuclear-retained transcripts?

In addition to CUTs, yeast also produce stable unannotated transcripts (SUTs). Whereas CUTs are only observable when the surveillance machinery is compromised, SUTs are readily detectable in wild-type cells [9]. The greater stability of SUTs might be explained by 'mRNA-like' processing [105], because they undergo 3′ cleavage and polyadenylation, with recruitment of Nab2 and/or Pab1 to the poly(A) tail. However, unlike most mRNAs, SUTs are not extensively bound by the export factor Mex67 or the cytoplasmic localization protein Hek2 [105]. These observations suggest SUTs are predominantly restricted to the nucleus, and perhaps targeted nonspecifically by nuclear RNA decay pathways. Consistently, SUTs are stabilized in strains lacking a functional nuclear exosome [32,130]. A subset of SUT-like RNAs are exported and strongly stabilized in the absence of the cytoplasmic exonuclease Xrn1, resulting in the designation of Xrn1-sensitive unstable transcripts (XUTs) [130].

SUT biogenesis appears homologous to nuclear lncRNAs in higher eukaryotes. Mammalian lncRNAs bear a striking resemblance to mRNAs; both classes are capped and polyadenylated, sometimes spliced, and similar in length [131]. However, lncRNAs usually lack the strong export and translation signals typical of most mRNAs, and are more frequently confined to the nucleus [132]. Only a small proportion of lncRNAs have defined functional roles, and few show any meaningful evolutionary conservation [119].

Like SUTs, nuclear lncRNAs are significantly less stable than their cytoplasmic mRNA counterparts [119]. The relative instability of SUTs and nuclear lncRNAs is commonly attributed to their inefficient processing and export, rather than specific recognition by quality control pathways. In general, species that take longer to exit the nucleus should be more susceptible to nuclear degradation. This 'kinetic competition' model has also been invoked to explain mRNA surveillance. In both yeast and humans, mistakes in splicing, 3′ end formation or mRNP packaging can inhibit nuclear export. In this way, some aberrant mRNAs may be degraded by default on account of their prolonged nuclear retention (figure 1).

7. Conclusion and perspectives

RNAs carry out diverse functions within cells, providing both the machinery for protein production and the information to programme its activity. In order to function in these distinct

activities, all RNAs undergo elaborate maturation processes within the nucleus. During these complex processing and assembly pathways, defects inevitably arise in some RNA transcripts and RNA–protein complexes, which must be systematically removed by quality control pathways. In this review, we proposed that the identification and degradation of defective RNAs, and enormous numbers of spurious transcripts, does not require recognition of specific 'bad' features. Rather the surveillance system will, by default, target almost all RNAs. Transcripts that undergo correct and timely maturation acquire protective features that help them evade the activities of the surveillance system. This model potentially explains how eukaryotes can tolerate the synthesis of huge numbers of diverse ncRNA transcripts, which are constantly cleared by rapid degradation. Moreover, it has been proposed, on thermodynamic grounds, that pervasive transcription is an inevitable feature of eukaryotic genomes [133]. The development of a surveillance ready transcription system may, therefore, have been a prerequisite for the evolution of the very large genomes found in many higher eukaryotes.

Although our understanding of these surveillance pathways is rapidly expanding, many questions remain unanswered. In particular, a quantitative description of RNA surveillance is needed. Discrimination between normal and defective transcripts must require an input of energy, in order to avoid violating the second law of thermodynamics. This energy input could be supplied by ATP-dependent RNA helicases, potentially explaining why helicases such as Mtr4 are critical components of essentially all surveillance pathways. However, a mechanistic understanding of the connection between ATP expenditure and RNA surveillance remains elusive. An additional avenue for future research is the role of RNA modifications such as m^6A and m^1A, which in recent years have been implicated in multiple aspects of the mRNA life cycle [134]. Nascent mRNAs must presumably be screened for the 'correct' combination of methylation marks, but how this is achieved is still unclear. In the future, the application of new techniques such as CRISPR, cross-linking-immunoprecipitation and single-RNA fluorescence microscopy will enable these questions and others to be addressed.

References

1. Hsin JP, Manley JL. 2012 The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137. (doi:10.1101/gad.200303.112)

2. Frenk S, Oxley D, Houseley J. 2014 The nuclear exosome is active and important during budding yeast meiosis. *PLoS ONE* **9**, e107648. (doi:10.1371/journal.pone.0107648)

3. Bitton DA *et al.* 2011 Programmed fluctuations in sense/antisense transcript ratios drive sexual differentiation in *S. pombe. Mol. Syst. Biol.* **7**, 559. (doi:10.1038/msb.2011.90)

4. Jiao X, Xiang S, Oh C, Martin CE, Tong L, Kiledjian M. 2010 Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* **467**, 608–611. (doi:10.1038/nature09338)

5. Hilleren P, McCarthy T, Rosbash M, Parker R, Jensen TH. 2001 Quality control of mRNA 3'-end processing is linked to the nuclear exosome. *Nature* **413**, 538–542. (doi:10.1038/35097110)

6. Saguez C, Schmid M, Olesen JR, Ghazy MA, Qu X, Poulsen MB, Nasser T, Moore C, Jensen TH. 2008 Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. *Mol. Cell* **31**, 91–103. (doi:10.1016/j.molcel.2008.04.030)

7. Porrua O, Libri D. 2015 Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* **16**, 190–202. (doi:10.1038/nrm3943)

8. 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. (doi:10.1038/nature07747)

9. Xu Z *et al.* 2009 Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**, 1033–1037. (doi:10.1038/nature07728)

10. Core LJ, Waterfall JJ, Lis JT. 2008 Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848. (doi:10.1126/science.1162228)

11. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. 2008 Divergent transcription from active promoters. *Science* **322**, 1849–1851. (doi:10.1126/science.1162253)

12. Ntini E *et al.* 2013 Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat. Struct. Mol. Biol.* **20**, 923–928. (doi:10.1038/nsmb.2640)

13. Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH. 2008 RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854. (doi:10.1126/science.1164096)

14. 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. (doi:10.1016/j.molcel.2006.07.024)

15. 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. (doi:10.1016/j.molcel.2006.07.029)

16. Wyers F *et al.* 2005 Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737. (doi:10.1016/j.cell.2005.04.030)

17. Davis CA, Ares Jr M. 2006 Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae. Proc. Natl Acad. Sci. USA* **103**, 3262–3267. (doi:10.1073/pnas.0507783103)

18. Stevens A, Poole TL. 1995 5'-exonuclease-2 of *Saccharomyces cerevisiae.* Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1. *J. Biol. Chem.* **270**, 16 063–16 069. (doi:10.1074/jbc.270.27.16063)

19. Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S. 2004 The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522. (doi:10.1038/nature03041)

20. Luo W, Johnson AW, Bentley DL. 2006 The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev.* **20**, 954–965. (doi:10.1101/gad.1409106)

21. Eaton JD, Davidson L, Bauer DLV, Natsume T, Kanemaki MT, West S. 2018 Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev.* **32**, 127–139. (doi:10.1101/gad.308528.117)

22. Jimeno-Gonzalez S, Haaning LL, Malagon F, Jensen TH. 2010 The yeast 5'–3' exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Mol. Cell* **37**, 580–587. (doi:10.1016/j.molcel.2010.01.019)

23. Xiang S, Cooper-Morgan A, Jiao X, Kiledjian M, Manley JL, Tong L. 2009 Structure and function of

the 5′ − >3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature* **458**, 784 – 788. (doi:10.1038/nature07731)

24. Johnson AW. 1997 Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. *Mol. Cell. Biol.* **17**, 6122 – 6130. (doi:10.1128/MCB.17.10.6122)

25. West S, Gromak N, Proudfoot NJ. 2004 Human 5′ − > 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522 – 525. (doi:10.1038/nature03035)

26. Brannan K *et al.* 2012 mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell* **46**, 311 – 324. (doi:10.1016/j.molcel.2012.03.006)

27. Chang JH, Jiao X, Chiba K, Oh C, Martin CE, Kiledjian M, Tong L. 2012 Dxo1 is a new type of eukaryotic enzyme with both decapping and 5′ – 3′ exoribonuclease activity. *Nat. Struct. Mol. Biol.* **19**, 1011 – 1017. (doi:10.1038/nsmb.2381)

28. Jiao X, Chang JH, Kilic T, Tong L, Kiledjian M. 2013 A mammalian pre-mRNA 5′ end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol. Cell* **50**, 104 – 115. (doi:10.1016/j.molcel.2013.02.017)

29. Celesnik H, Deana A, Belasco JG. 2007 Initiation of RNA decay in *Escherichia coli* by 5′ pyrophosphate removal. *Mol. Cell* **27**, 79 – 90. (doi:10.1016/j.molcel.2007.05.038)

30. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. 1999 Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399 – 5410. (doi:10.1093/emboj/18.19.5399)

31. Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D. 1997 The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′ − >5′ exoribonucleases. *Cell* **91**, 457 – 466. (doi:10.1016/S0092-8674(00)80432-8)

32. Gudipati RK, Xu Z, Lebreton A, Seraphin B, Steinmetz LM, Jacquier A, Libri D. 2012 Extensive degradation of RNA precursors by the exosome in wild-type cells. *Mol. Cell* **48**, 409 – 421. (doi:10.1016/j.molcel.2012.08.018)

33. Schneider C, Kudla G, Wlotzka W, Tuck A, Tollervey D. 2012 Transcriptome-wide analysis of exosome targets. *Mol. Cell* **48**, 422 – 433. (doi:10.1016/j.molcel.2012.08.013)

34. Chekanova JA *et al.* 2007 Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* **131**, 1340 – 1353. (doi:10.1016/j.cell.2007.10.056)

35. Szczepinska T, Kalisiak K, Tomecki R, Labno A, Borowski LS, Kulinski TM, Adamska D, Kosinska J, Dziembowski A. 2015 DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res.* **25**, 1622 – 1633. (doi:10.1101/gr.189597.115)

36. Bousquet-Antonelli C, Presutti C, Tollervey D. 2000 Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102**, 765 – 775. (doi:10.1016/S0092-8674(00)00065-9)

37. Kilchert C, Wittmann S, Vasiljeva L. 2016 The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* **17**, 227 – 239. (doi:10.1038/nrm.2015.15)

38. Schneider C, Tollervey D. 2013 Threading the barrel of the RNA exosome. *Trends Biochem. Sci.* **38**, 485 – 493. (doi:10.1016/j.tibs.2013.06.013)

39. Chlebowski A, Lubas M, Jensen TH, Dziembowski A. 2013 RNA decay machines: the exosome. *Biochim. Biophys. Acta* **1829**, 552 – 560. (doi:10.1016/j.bbagrm.2013.01.006)

40. Makino DL, Baumgartner M, Conti E. 2013 Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* **495**, 70 – 75. (doi:10.1038/nature11870)

41. Wasmuth EV, Lima CD. 2012 Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. *Mol. Cell* **48**, 133 – 144. (doi:10.1016/j.molcel.2012.07.012)

42. Schuch B, Feigenbutz M, Makino DL, Falk S, Basquin C, Mitchell P, Conti E. 2014 The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. *EMBO J.* **33**, 2829 – 2846. (doi:10.15252/embj.201488757)

43. Mitchell P, Petfalski E, Houalla R, Podtelejnikov A, Mann M, Tollervey D. 2003 Rrp47p is an exosome-associated protein required for the 3′ processing of stable RNAs. *Mol. Cell. Biol.* **23**, 6982 – 6992. (doi:10.1128/MCB.23.19.6982-6992.2003)

44. Schilders G, van Dijk E, Pruijn GJ. 2007 C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. *Nucleic Acids Res.* **35**, 2564 – 2572. (doi:10.1093/nar/gkm082)

45. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005 RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**, 713 – 724. (doi:10.1016/j.cell.2005.04.029)

46. Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith GK, Keller W. 2005 A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.* **3**, e189. (doi:10.1371/journal.pbio.0030189)

47. Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. 2004 Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes Dev.* **18**, 1227 – 1240. (doi:10.1101/gad.1183804)

48. de la Cruz J, Kressler D, Tollervey D, Linder P. 1998 Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3′ end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J.* **17**, 1128 – 1140. (doi:10.1093/emboj/17.4.1128)

49. Thoms M, Thomson E, Bassler J, Gnadig M, Griesel S, Hurt E. 2015 The exosome is recruited to RNA substrates through specific adaptor proteins. *Cell* **162**, 1029 – 1038. (doi:10.1016/j.cell.2015.07.060)

50. Dez C, Houseley J, Tollervey D. 2006 Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of *Saccharomyces cerevisiae*. *EMBO J.* **25**, 1534 – 1546. (doi:10.1038/sj.emboj.7601035)

51. Wery M, Ruidant S, Schillewaert S, Lepore N, Lafontaine DL. 2009 The nuclear poly(A) polymerase and exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance. *RNA* **15**, 406 – 419. (doi:10.1261/rna.1402709)

52. Houseley J, Tollervey D. 2006 Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO Rep.* **7**, 205 – 211. (doi:10.1038/sj.embor.7400612)

53. Schulz D, Schwalb B, Kiesel A, Baejen C, Torkler P, Gagneur J, Soeding J, Cramer P. 2013 Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* **155**, 1075 – 1087. (doi:10.1016/j.cell.2013.10.024)

54. Wlotzka W, Kudla G, Granneman S, Tollervey D. 2011 The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* **30**, 1790 – 1803. (doi:10.1038/emboj.2011.97)

55. Creamer TJ, Darby MM, Jamonnak N, Schaughency P, Hao H, Wheelan SJ, Corden JL. 2011 Transcriptome-wide binding sites for components of the *Saccharomyces cerevisiae* non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS Genet.* **7**, e1002329. (doi:10.1371/journal.pgen.1002329)

56. Webb S, Hector RD, Kudla G, Granneman S. 2014 PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol.* **15**, R8. (doi:10.1186/gb-2014-15-1-r8)

57. 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. (doi:10.1038/nsmb.1460)

58. 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. (doi:10.1038/nsmb.1468)

59. Holmes RK *et al.* 2015 Loss of the yeast SR protein Npl3 alters gene expression due to transcription readthrough. *PLoS Genet.* **11**, e1005735. (doi:10.1371/journal.pgen.1005735)

60. Tudek A *et al.* 2014 Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Mol. Cell* **55**, 467 – 481. (doi:10.1016/j.molcel.2014.05.031)

61. Fasken MB, Laribee RN, Corbett AH. 2015 Nab3 facilitates the function of the TRAMP complex in RNA processing via recruitment of Rrp6 independent of Nrd1. *PLoS Genet.* **11**, e1005044. (doi:10.1371/journal.pgen.1005044)

62. Vasiljeva L, Buratowski S. 2006 Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. *Mol. Cell* **21**, 239 – 248. (doi:10.1016/j.molcel.2005.11.028)

63. Terzi N, Churchman LS, Vasiljeva L, Weissman J, Buratowski S. 2011 H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol. Cell. Biol.* **31**, 3569 – 3583. (doi:10.1128/MCB.05590-11)

64. 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. (doi:10.1038/35095090)

65. Schmidt K, Butler JS. 2013 Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. *Wiley Interdiscip. Rev. RNA* **4**, 217–231. (doi:10.1002/wrna.1155)

66. Lubas M, Christensen MS, Kristiansen MS, Domanski M, Falkenby LG, Lykke-Andersen S, Andersen JS, Dziembowski A, Jensen TH. 2011 Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637. (doi:10.1016/j.molcel.2011.06.028)

67. Hrossova D, Sikorsky T, Potesil D, Bartosovic M, Pasulka J, Zdrahal Z, Stefl R, Vanacova S. 2015 RBM7 subunit of the NEXT complex binds U-rich sequences and targets 3'-end extended forms of snRNAs. *Nucleic Acids Res.* **43**, 4236–4248. (doi:10.1093/nar/gkv240)

68. Andersen PR *et al.* 2013 The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.* **20**, 1367–1376. (doi:10.1038/nsmb.2703)

69. Lubas M, Andersen PR, Schein A, Dziembowski A, Kudla G, Jensen TH. 2015 The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.* **10**, 178–192. (doi:10.1016/j.celrep.2014.12.026)

70. Nudler E. 2012 RNA polymerase backtracking in gene regulation and genome instability. *Cell* **149**, 1438–1445. (doi:10.1016/j.cell.2012.06.003)

71. Izban MG, Luse DS. 1992 The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'–5' direction in the presence of elongation factor SII. *Genes Dev.* **6**, 1342–1356. (doi:10.1101/gad.6.7.1342)

72. Lemay JF, Larochelle M, Marguerat S, Atkinson S, Bahler J, Bachand F. 2014 The RNA exosome promotes transcription termination of backtracked RNA polymerase II. *Nat. Struct. Mol. Biol.* **21**, 919–926. (doi:10.1038/nsmb.2893)

73. Gromak N, West S, Proudfoot NJ. 2006 Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996. (doi:10.1128/MCB.26.10.3986-3996.2006)

74. Glover-Cutter K, Kim S, Espinosa J, Bentley DL. 2008 RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.* **15**, 71–78. (doi:10.1038/nsmb1352)

75. Fox MJ, Gao H, Smith-Kinnaman WR, Liu Y, Mosley AL. 2015 The exosome component Rrp6 is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway. *PLoS Genet.* **11**, e1004999. (doi:10.1371/journal.pgen.1004999)

76. Schwalb B, Michel M, Zacher B, Fruhauf K, Demel C, Tresch A, Gagneur J, Cramer P. 2016 TT-seq maps the human transient transcriptome. *Science* **352**, 1225–1228. (doi:10.1126/science.aad9841)

77. Kilchert C, Wittmann S, Passoni M, Shah S, Granneman S, Vasiljeva L. 2015 Regulation of mRNA levels by decay-promoting introns that recruit the exosome specificity factor Mmi1. *Cell Rep.* **13**, 2504–2515. (doi:10.1016/j.celrep.2015.11.026)

78. Eckmann CR, Rammelt C, Wahle E. 2011 Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2**, 348–361. (doi:10.1002/wrna.56)

79. Schmid M, Olszewski P, Pelechano V, Gupta I, Steinmetz LM, Jensen TH. 2015 The nuclear polyA-binding protein Nab2p Is essential for mRNA production. *Cell Rep.* **12**, 128–139. (doi:10.1016/j.celrep.2015.06.008)

80. Colin J, Candelli T, Porrua O, Boulay J, Zhu C, Lacroute F, Steinmetz LM, Libri D. 2014 Roadblock termination by reb1p restricts cryptic and readthrough transcription. *Mol. Cell* **56**, 667–680. (doi:10.1016/j.molcel.2014.10.026)

81. Rondon AG, Mischo HE, Kawauchi J, Proudfoot NJ. 2009 Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Mol. Cell* **36**, 88–98. (doi:10.1016/j.molcel.2009.07.028)

82. Ghazal G, Gagnon J, Jacques PE, Landry JR, Robert F, Elela SA. 2009 Yeast RNase III triggers polyadenylation-independent transcription termination. *Mol. Cell* **36**, 99–109. (doi:10.1016/j.molcel.2009.07.029)

83. Wilusz JE. 2016 Long noncoding RNAs: re-writing dogmas of RNA processing and stability. *Biochim. Biophys. Acta* **1859**, 128–138. (doi:10.1016/j.bbagrm.2015.06.003)

84. Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. 2012 Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs. *Proc. Natl Acad. Sci. USA* **109**, 19 202–19 207. (doi:10.1073/pnas.1217338109)

85. Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012 A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **26**, 2392–2407. (doi:10.1101/gad.204438.112)

86. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. 2010 Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* **330**, 1244–1247. (doi:10.1126/science.1195858)

87. Conrad NK, Mili S, Marshall EL, Shu MD, Steitz JA. 2006 Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. *Mol. Cell* **24**, 943–953. (doi:10.1016/j.molcel.2006.10.029)

88. Bresson SM, Conrad NK. 2013 The human nuclear poly(a)-binding protein promotes RNA hyperadenylation and decay. *PLoS Genet.* **9**, e1003893. (doi:10.1371/journal.pgen.1003893)

89. Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA. 2012 Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Rep.* **2**, 26–32. (doi:10.1016/j.celrep.2012.05.020)

90. Tycowski KT, Shu MD, Steitz JA. 2016 Myriad triple-helix-forming structures in the transposable element RNAs of plants and fungi. *Cell Rep.* **15**, 1266–1276. (doi:10.1016/j.celrep.2016.04.010)

91. Marzluff WF, Wagner EJ, Duronio RJ. 2008 Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* **9**, 843–854. (doi:10.1038/nrg2438)

92. Pannone BK, Xue D, Wolin SL. 1998 A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.* **17**, 7442–7453. (doi:10.1093/emboj/17.24.7442)

93. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. 1999 A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'- end of U6 snRNA, thereby facilitating U4/U6 duplex formation *in vitro*. *EMBO J.* **18**, 5789–5802. (doi:10.1093/emboj/18.20.5789)

94. Fernandez CF, Pannone BK, Chen X, Fuchs G, Wolin SL. 2004 An Lsm2-Lsm7 complex in *Saccharomyces cerevisiae* associates with the small nucleolar RNA snR5. *Mol. Biol. Cell* **15**, 2842–2852. (doi:10.1091/mbc.E04-02-0116)

95. Kufel J, Allmang C, Verdone L, Beggs J, Tollervey D. 2003 A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic Acids Res.* **31**, 6788–6797. (doi:10.1093/nar/gkg904)

96. Coy S, Volanakis A, Shah S, Vasiljeva L. 2013 The Sm complex is required for the processing of non-coding RNAs by the exosome. *PLoS ONE* **8**, e65606. (doi:10.1371/journal.pone.0065606)

97. Pefanis E *et al.* 2015 RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* **161**, 774–789. (doi:10.1016/j.cell.2015.04.034)

98. Wahba L, Amon JD, Koshland D, Vuica-Ross M. 2011 RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* **44**, 978–988. (doi:10.1016/j.molcel.2011.10.017)

99. Hage AE, Webb S, Kerr A, Tollervey D. 2014 Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* **10**, e1004716. (doi:10.1371/journal.pgen.1004716)

100. Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN, Basu U. 2014 Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* **514**, 389–393. (doi:10.1038/nature13580)

101. Schmid M, Poulsen MB, Olszewski P, Pelechano V, Saguez C, Gupta I, Steinmetz LM, Moore C, Jensen TH. 2012 Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. *Mol. Cell* **47**, 267–280. (doi:10.1016/j.molcel.2012.05.005)

102. Soucek S *et al.* 2016 The evolutionarily-conserved polyadenosine RNA binding protein, Nab2, cooperates with splicing machinery to regulate the fate of pre-mRNA. *Mol. Cell. Biol.* **36**, 2697–2714. (doi:10.1128/MCB.00402-16)

103. 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. (doi:10.1261/rna.709609)

104. Roth KM, Wolf MK, Rossi M, Butler JS. 2005 The nuclear exosome contributes to autogenous control of NAB2 mRNA levels. *Mol. Cell. Biol.* **25**, 1577–1585. (doi:10.1128/MCB.25.5.1577-1585.2005)

105. Tuck AC, Tollervey D. 2013 A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* **154**, 996–1009. (doi:10.1016/j.cell.2013.07.047)

106. Beaulieu YB, Kleinman CL, Landry-Voyer AM, Majewski J, Bachand F. 2012 Polyadenylation-dependent control of long noncoding RNA expression by the poly(A)-binding protein nuclear 1. *PLoS Genet.* **8**, e1003078. (doi:10.1371/journal.pgen.1003078)

107. Li W *et al.* 2015 Systematic profiling of poly(A)+ transcripts modulated by core 3′ end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet.* **11**, e1005166. (doi:10.1371/journal.pgen.1005166)

108. Meola N *et al.* 2016 Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell* **64**, 520–533. (doi:10.1016/j.molcel.2016.09.025)

109. Bresson SM, Hunter OV, Hunter AC, Conrad NK. 2015 Canonical Poly(A) polymerase activity promotes the decay of a wide variety of mammalian nuclear RNAs. *PLoS Genet.* **11**, e1005610. (doi:10.1371/journal.pgen.1005610)

110. Kuhn U, Gundel M, Knoth A, Kerwitz Y, Rudel S, Wahle E. 2009 Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J. Biol. Chem.* **284**, 22 803–22 814. (doi:10.1074/jbc.M109.018226)

111. Ogami K, Richard P, Chen Y, Hoque M, Li W, Moresco JJ, Yates JR, Tian B, Manley JL. 2017 An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* **31**, 1257–1271. (doi:10.1101/gad.302604.117)

112. Lee NN *et al.* 2013 Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. *Cell* **155**, 1061–1074. (doi:10.1016/j.cell.2013.10.027)

113. Lemieux C, Marguerat S, Lafontaine J, Barbezier N, Bahler J, Bachand F. 2011 A pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. *Mol. Cell* **44**, 108–119. (doi:10.1016/j.molcel.2011.06.035)

114. Yamanaka S, Yamashita A, Harigaya Y, Iwata R, Yamamoto M. 2010 Importance of polyadenylation in the selective elimination of meiotic mRNAs in growing *S. pombe* cells. *EMBO J.* **29**, 2173–2181. (doi:10.1038/emboj.2010.108)

115. Sugiyama T, Sugioka-Sugiyama R. 2011 Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *EMBO J.* **30**, 1027–1039. (doi:10.1038/emboj.2011.32)

116. Lemay JF, D'Amours A, Lemieux C, Lackner DH, St-Sauveur VG, Bahler J, Bachand F. 2010 The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol. Cell* **37**, 34–45. (doi:10.1016/j.molcel.2009.12.019)

117. Zhou Y, Zhu J, Schermann G, Ohle C, Bendrin K, Sugioka-Sugiyama R, Sugiyama T, Fischer T. 2015 The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. *Nat. Commun.* **6**, 7050. (doi:10.1038/ncomms8050)

118. Kessler MM, Henry MF, Shen E, Zhao J, Gross S, Silver PA, Moore CL. 1997 Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3′-end formation in yeast. *Genes Dev.* **11**, 2545–2556. (doi:10.1101/gad.11.19.2545)

119. Mukherjee N, Calviello L, Hirsekorn A, de Pretis S, Pelizzola M, Ohler U. 2017 Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat. Struct. Mol. Biol.* **24**, 86–96. (doi:10.1038/nsmb.3325)

120. Kapranov P *et al.* 2007 RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484–1488. (doi:10.1126/science.1138341)

121. Jin Y, Eser U, Struhl K, Churchman LS. 2017 The ground state and evolution of promoter region directionality. *Cell* **170**, 889–898 e10. (doi:10.1016/j.cell.2017.07.006)

122. Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. 2012 CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* **336**, 1723–1725. (doi:10.1126/science.1219651)

123. Milligan L, Huynh-Thu VA, Delan-Forino C, Tuck A, Petfalski E, Lombrana R, Sanguinetti G, Kudla G, Tollervey D. 2016 Strand-specific, high-resolution mapping of modified RNA polymerase II. *Mol. Syst. Biol.* **12**, 874. (doi:10.15252/msb.20166869)

124. Lunde BM *et al.* 2010 Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* **17**, 1195–1201. (doi:10.1038/nsmb.1893)

125. Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. 2013 Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363. (doi:10.1038/nature12349)

126. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. 2010 U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668. (doi:10.1038/nature09479)

127. Chiu AC, Suzuki HI, Wu X, Mahat DB, Kriz AJ, Sharp PA. 2018 Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. *Mol. Cell* **69**, 648–663.37. (doi:10.1016/j.molcel.2018.01.006)

128. Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R. 2006 Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell* **127**, 1389–1400. (doi:10.1016/j.cell.2006.10.044)

129. Fan J *et al.* 2017 Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J.* **36**, 2870–2886. (doi:10.15252/embj.201696139)

130. Marquardt S, Hazelbaker DZ, Buratowski S. 2011 Distinct RNA degradation pathways and 3′ extensions of yeast non-coding RNA species. *Transcription* **2**, 145–154. (doi:10.4161/trns.2.3.16298)

131. Guttman M *et al.* 2009 Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227. (doi:10.1038/nature07672)

132. Quinn JJ, Chang HY. 2016 Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17**, 47–62. (doi:10.1038/nrg.2015.10)

133. Struhl K. 2007 Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* **14**, 103–105. (doi:10.1038/nsmb0207-103)

134. Roundtree IA, Evans ME, Pan T, He C. 2017 Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200. (doi:10.1016/j.cell.2017.05.045)