Attenuation of Eukaryotic Protein-Coding Gene Expression via Premature Transcription Termination

DEIRDRE C. TATOMER AND JEREMY E. WILUSZ

Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania 19104, USA

Correspondence: wilusz@pennmedicine.upenn.edu

A complex network of RNA transcripts is generated from eukaryotic genomes, many of which are processed in unexpected ways. Here, we highlight how premature transcription termination events at protein-coding gene loci can simultaneously lead to the generation of short RNAs and attenuate production of full-length mRNA transcripts. We recently showed that the Integrator (Int) complex can be selectively recruited to protein-coding gene loci, including *Drosophila* metallothionein A (MtnA), where the IntS11 RNA endonuclease cleaves nascent transcripts near their 5' ends. Such premature termination events catalyzed by Integrator can repress the expression of some full-length mRNAs by more than 100-fold. Transcription at small nuclear RNA (snRNA) loci is likewise terminated by Integrator cleavage, but protein-coding and snRNA gene loci have notably distinct dependencies on Integrator subunits. Additional mechanisms that attenuate eukaryotic gene outputs via premature termination have been discovered, including by the cleavage and polyadenylation machinery in a manner controlled by U1 snRNP. These mechanisms appear to function broadly across the transcriptome. This suggests that synthesis of full-length transcripts is not always the default option and that premature termination events can lead to a variety of transcripts, some of which may have important and unexpected biological functions.

It is well established that nascent transcripts must be extensively processed in order to generate a mature RNA that is stable and functional. For transcripts derived from protein-coding genes, they typically must be capped at their 5' ends, spliced to remove intronic sequences, and modified at their 3' ends by the addition of a poly(A) tail. Each of these canonical pre-mRNA processing steps is extensively regulated, ensuring quality control as well as enabling the generation of a diverse set of functional RNAs (for reviews, see Moore and Proudfoot 2009; Braunschweig et al. 2013; Tian and Manley 2017; Hentze et al. 2018). Besides well-studied transcripts that undergo all of the canonical pre-mRNA processing steps (Fig. 1A), a growing number of transcripts that lack a standard 5' cap, a poly(A) tail, or both of these terminal structures have been identified, with some accumulating to high levels and having key cellular roles (for reviews, see Wilusz and Spector 2010; Zhang et al. 2014; Wilusz 2016; Kiledjian 2018). For example, metazoan replicative histone mRNAs are efficiently translated despite ending in a stem-loop structure (Fig. 1B; for review, see Marzluff et al. 2008); the MALAT1 and MEN β (also known as NEAT1_2) nuclear-retained long noncoding RNAs have 3' terminal triple helices (Wilusz et al. 2008, 2012; Brown et al. 2012, 2014) and play key roles in cancer metastasis (Ji et al. 2003; Arun et al. 2016) and nuclear paraspeckle formation (Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009; Mao et al. 2011), respectively; and circular RNAs have covalently linked ends and, in some cases, regulate microRNA activity, the immune system, or other cellular pathways (for reviews, see Li et al.

2018; Wilusz 2018; Kristensen et al. 2019; Patop et al. 2019).

Because nonpolyadenylated RNAs and circular RNAs are structurally distinct from canonical mRNAs, they are subjected to different biogenesis and post-transcriptional control mechanisms as well as likely bound by unique factors. A major goal of our laboratory has thus been to understand how the fates of noncanonical RNAs are controlled. This has led us in a variety of scientific directions and led to a number of unexpected findings, including mechanisms that control (i) circular RNA levels and localization (Liang and Wilusz 2014; Kramer et al. 2015; Liang et al. 2017; Huang et al. 2018), (ii) non-AUG translation (Kearse et al. 2019), and (iii) transcription elongation (Tatomer et al. 2019), which will be the focus of this manuscript. We will summarize how a high-throughput RNAi screening effort revealed that the Integrator (Int) complex is a potent inhibitor of the transcription of many protein-coding genes. This is because the IntS11 RNA endonuclease cleaves nascent transcripts and catalyzes premature transcription termination. Additional mechanisms that attenuate eukaryotic gene outputs via premature termination have been discovered and may likewise be widespread across the transcriptome (for a review, see Kamieniarz-Gdula and Proudfoot 2019). This suggests that synthesis of full-length transcripts is likely not the default option. Instead, premature cleavage events by the Integrator complex, the cleavage and polyadenylation (CPA) machinery, and likely other endonucleases need to be actively suppressed for full-length transcripts to be generated.

Figure 1. RNA 3'-end processing mechanisms. Proper processing of the RNA 3' end is critical for the generation of a mature, functional transcript. (A) For most protein-coding genes, the cleavage and polyadenylation machinery are recruited to a polyadenylation signal (PAS) with help from a downstream sequence element (DSE). The Cpsf73 endonuclease (denoted 73) cleaves the nascent transcript, which is then modified by the addition of a poly(A) tail. (B) At replication-dependent histone genes, nascent transcripts are cleaved by Cpsf73 (which is part of the histone cleavage complex, HCC) between a stem-loop structure and a purine-rich histone downstream element (HDE). During histone mRNA 3'-end processing, the HDE interacts with the U7 snRNP and the stem-loop is recognized by stem-loop binding protein (SLBP). Cleavage is not followed by addition of a poly(A) tail. (C) Nascent snRNA transcripts are recognized by the Integrator complex and cleaved upstream of a 3' box sequence by the IntS11 endonuclease (denoted 11), which is homologous to Cpsf73 and is part of a core cleavage complex within Integrator.

RNAi SCREENING REVEALED THE INTEGRATOR COMPLEX AS A POTENT NEGATIVE REGULATOR OF *DROSOPHILA* MtnA TRANSCRIPTION

To identify novel factors that control RNA fates, we worked with Sara Cherry's group to set up genome-scale RNAi screens in *Drosophila* cells. In particular, we examined the outputs of inducible reporter mRNAs that ended in distinct 3' terminal sequences—for example, eGFP or mCherry mRNAs ending in a poly(A) tail, the histone stem-loop structure, or the MALAT1 triple helix (Fig. 2A). By quantitating the levels of the encoded fluorescent proteins using automated microscopy, we reasoned that this approach should enable the identification of (i) general factors that regulate all of the reporter RNAs (e.g., general transcription and translation control factors), (ii) factors that regulate all of the nonpolyadenylated RNAs, and (iii) factors that are specific for particular 3' ends (e.g., stem-loop binding protein for the histone stem-loop [Wang et al. 1996]). Among other observations, these efforts surprisingly revealed that nearly all

14 subunits of the Int complex (for a review, see Baillat and Wagner 2015) were among the most potent negative regulators of the reporters (Fig. 2B,C; Tatomer et al. 2019). When many Integrator subunits were individually depleted, we observed increased levels of the reporter RNAs and encoded reporter proteins, and these effects were observed regardless of the open reading frame or 3' end processing mechanism that was present downstream from the inducible metallothionein A (MtnA) promoter (Fig. 2B,C). This suggested that the Integrator complex inhibits the transcriptional output of the MtnA promoter.

The >1 MDa Integrator complex is conserved across metazoans, interacts with RNA polymerase II (RNAPII), and consists of 14 subunits, but few of the subunits have identifiable paralogs within eukaryotic genomes (Baillat et al. 2005; Egloff et al. 2010; Baillat and Wagner 2015). Strikingly, many Integrator subunits are devoid of known protein domains, with the most common domains being α -helical repeats (e.g., HEAT, ARM, TPR, or VWA domains) that may function as protein-protein interaction surfaces. Notable exceptions are Integrator subunits 11 (IntS11) and 9 (IntS9) that are homologous to Cpsf73

TRANSCRIPTION ATTENUATION VIA PREMATURE TERMINATION

3

Figure 2. The Integrator complex inhibits expression of reporter mRNAs generated from the MtnA promoter. (A) A set of reporter plasmids was generated that encode fluorescent proteins and are processed by distinct RNA 3'-end processing mechanisms. All reporters are driven by the *Drosophila* MtnA promoter, which is rapidly induced when the intracellular concentration of heavy metals (e.g., copper or cadmium) is increased. The SV40 polyadenylation signal enables generation of an mRNA ending in a poly(A) tail, the histone mRNA processing signals enable generation of an mRNA ending in the histone stem-loop, and the self-cleaving hammerhead ribozyme (HhRz) enables generation of nonpolyadenylated reporter RNAs with defined 3' ends. Insertion of the MALAT1 3' end upstream of the HhRz sequence enables generation of an mRNA ending in a triple helix. (B,C) *Drosophila* DL1 cells stably maintaining an eGFP reporter ending in the SV40 polyadenylation signal or the MALAT1 triple helix were treated with double-stranded RNAs (dsRNAs) for 3 d to induce RNAi and depletion of the indicated Integrator subunits. CuSO₄ was added for the last 6 h. (B) Representative images of eGFP and DNA (Hoechst 33342). (C) The integrated eGFP intensity (amount of eGFP signal in each well divided by the number of cells) was quantified. Integrator depletion resulted in similar effects on eGFP protein levels regardless of the mRNA 3'-end processing mechanism.

and Cpsf100, which bind one another and function as a zinc-dependent endonuclease that is critical for cleaving the 3' ends of mRNAs prior to the addition of the poly(A) tail (Fig. 1; Baillat et al. 2005; Shi and Manley 2015). The IntS4, IntS9, and IntS11 proteins interact with one another (Fig. 1C; Wu et al. 2017; Albrecht et al. 2018) and, like the Cpsf73, Cpsf100, Symplekin complex (Fig. 1A; Sullivan et al. 2009), can catalyze cleavage of RNA. In particular, IntS11 has been well-established to cleave the 3' ends of nascent small nuclear RNA (snRNA) transcripts, thereby releasing the processed snRNA from RNAPII so that it can

function in pre-mRNA splicing as part of the spliceosome (Fig. 1C; Baillat et al. 2005). Until fairly recently, snRNA cleavage was the only known function for Integrator, but recent work has suggested that this complex may also bind to a subset of protein-coding, noncoding, and enhancer loci to control their expression (Cazalla et al. 2011; Gardini et al. 2014; Stadelmayer et al. 2014; Lai et al. 2015; Skaar et al. 2015; Xie et al. 2015; Barbieri et al. 2018; Rubtsova et al. 2019).

We, therefore, examined binding of Integrator subunits to the endogenous MtnA locus using chromatin immuno-

Figure 3. The Integrator complex regulates the outputs of endogenous snRNA and MtnA genes. *Drosophila* DL1 cells were treated with dsRNAs for 3 d to induce RNAi and depletion of the indicated Integrator subunits. RT-qPCR was used to quantify expression of endogenous MtnA mRNA after 14 h of CuSO₄ treatment. Northern blotting was used to quantify readthrough transcription downstream from the U4:39B snRNA. Data are shown as mean $\pm$ SD, $N \geq 3$. (*) $P < 0.05$.

precipitation (ChIP)-quantitative polymerase chain reaction (qPCR) in collaboration with Eric J. Wagner's group. The MtnA locus encodes a metal chelator, and its promoter is rapidly induced when the intracellular concentration of heavy metals (e.g., copper or cadmium) is increased (Günther et al. 2012). Much more significant binding of Integrator subunits to the 5' end of the MtnA locus was noted during copper stress (the same conditions used for the genome-scale RNAi screens) as compared to cadmium stress (Tatomer et al. 2019). Consistent with these ChIP-qPCR results, depletion of Integrator subunits resulted in increased levels of MtnA pre-mRNA and mRNA during copper (Fig. 3) but not cadmium stress. Integrator can thus display context-specific binding and regulation, including helping with homeostatic control of intracellular heavy metal levels via direct control of MtnA gene expression.

THE INTEGRATOR COMPLEX CLEAVES NASCENT MtnA TRANSCRIPTS TO TRIGGER PREMATURE TRANSCRIPTION TERMINATION

Individual depletion of many Integrator subunits, including noncatalytic subunits that currently lack a known function, resulted in increased MtnA expression (Fig. 3; Tatomer et al. 2019). Nevertheless, to reveal the underlying molecular mechanism, we first examined whether the well-established RNA endonuclease activity of IntS11 is required for the Integrator complex to limit the output from the MtnA promoter. Endogenous IntS11 was depleted from *Drosophila* cells using RNAi coupled to expression of either a wild-type or catalytically dead (E203Q) RNAi-resistant IntS11 transgene. These knockdown-rescue experiments revealed a clear requirement for the IntS11 RNA endonuclease activity (Tatomer et al. 2019) and sparked a search for where Integrator may be cleaving RNAs derived from the MtnA locus.

We reasoned that the RNA products of Integrator cleavage may be unstable because of the lack of a poly(A) tail and further noted that many RNA exosome core components and cofactors (including Rrp40 and Mtr4) had scored as positive regulators of the MtnA promoter in the genome-scale RNAi screens. We thus hypothesized that Integrator may cleave nascent MtnA transcripts to prematurely terminate transcription and that these cleaved transcripts are then targeted for rapid degradation by the RNA exosome. To test this model, the exosome-associated RNA helicase Mtr4 (Lubas et al. 2011) was depleted from cells during copper stress (Fig. 4, lane 4). This resulted in a reduction in full-length MtnA mRNA expression that was coupled to increased expression of several small RNAs, including prominent transcripts with lengths of ~85 and ~110 nt (Fig. 4, lane 4; Tatomer et al. 2019). These small RNAs (i) were capped at their 5' ends, (ii) had

Figure 4. The Integrator complex catalyzes premature transcription termination by cleaving nascent MtnA RNAs. Northern blotting was used to analyze RNAs generated from the endogenous MtnA locus in *Drosophila* DL1 cells treated with the indicated dsRNAs for 3 d and CuSO₄ for the last 6 h. Full-length MtnA mRNA (black arrow) and Integrator-dependent small RNAs (orange) are indicated. *, Nonspecific band.

Figure 5. Model for Integrator-dependent premature transcription termination. After transcription initiation, RNA polymerase II can transition into productive elongation to generate the mature mRNA. Alternatively, the IntS11 endonuclease within the Integrator complex can cleave the nascent RNA, enabling transcription termination and degradation of the short RNA by the RNA exosome.

the same transcription start site (TSS) as full-length MtnA mRNA and a similar requirement for the MTF-1 transcription factor (Fig. 4, lane 6), as well as (iii) had detectable oligoadenylation, a mark known to facilitate RNA degradation by the RNA exosome. These features are all consistent with premature transcription termination products. Upon then co-depleting Mtr4 and Integrator subunits, these small RNAs were eliminated and full-length MtnA mRNA expression was restored (Fig. 4, lanes 8 and 10). Generation of the MtnA small RNAs is thus dependent on the presence of Integrator and, in particular, catalytically active IntS11.

In total, these data indicate that the Integrator complex can be recruited to the MtnA locus, where it cleaves nascent RNAs and facilitates premature transcription termination (Fig. 5; Tatomer et al. 2019). Integrator cleavage limits production of full-length MtnA mRNA, likely as a homeostatic mechanism to fine-tune copper levels in cells. This is particularly important as copper is required for the function of a subset of enzymes and must be maintained in a narrow concentration range (for a review, see Festa and Thiele 2011).

THE INTEGRATOR COMPLEX CATALYZES PREMATURE TRANSCRIPTION TERMINATION AT MANY PROTEIN-CODING GENES

Given that Integrator cleavage potentially attenuates MtnA transcription, we next addressed whether additional protein-coding genes are similarly regulated. The IntS9 subunit was depleted from *Drosophila* cells (in copper stress conditions) and RNA-seq was used to identify mRNA expression changes across the transcriptome. This revealed 409 and 49 genes that were up- and down-regulated, respectively, upon IntS9 depletion (fold change >1.5 and $P < 0.001$) (Tatomer et al. 2019). Integrator thus predominately inhibits protein-coding gene expression in *Drosophila* cells, which is in contrast to previous reports in other systems that suggested that Integrator predominantly stimulates protein-coding gene expression (Gardini et al. 2014; Stadelmayer et al. 2014). Furthermore, we find that Integrator can function as a very potent inhibitor of gene expression, with some genes being repressed by

more than 25-fold by Integrator (e.g., *Pepck1* and *Hml* are repressed by ~100- and 35-fold, respectively), including in unstressed cells.

To validate these results, seven mRNAs that had differing magnitudes of fold change upon IntS9 depletion were selected for further analyses, which included reverse transcription (RT)-qPCR, ChIP-qPCR, and northern blotting to identify small RNAs (Tatomer et al. 2019). The results at these loci largely mirrored what was observed at the MtnA locus: (i) expression of the pre-mRNAs increased upon depletion of many Integrator subunits (including noncatalytic subunits), indicative of a transcriptional effect, (ii) Integrator subunits were bound at the 5' ends of these loci, but not at loci whose expression was unchanged upon Integrator depletion, (iii) the IntS11 endonuclease was required for regulation of these loci, and (iv) prematurely terminated small RNAs could be detected using northern blotting. Interestingly, the prematurely terminated small RNAs were of defined lengths and often 50–110 nt, roughly mirroring the sizes of cleavage products observed at the MtnA locus.

At snRNA gene loci, a conserved but relatively degenerate 3' box sequence is required for Integrator cleavage (Fig. 1C; Hernandez 1985), but similar sequences are not immediately obvious at any of the protein-coding transcripts we have examined in detail. Introducing deletions into the MtnA 5' UTR did not alter the cleavage product sizes (Tatomer et al. 2019), suggesting that Integrator may cleave nascent mRNAs at a set distance from the TSS in a manner independent of local DNA/RNA sequence, perhaps at positions of RNA polymerase II pausing/stalling or nucleosomes. Indeed, collaborative work with the laboratories of Karen Adelman and Eric J. Wagner suggests a role for RNA polymerase II pausing/stalling in dictating Integrator cleavage sites (Elrod et al. 2019). It will be very informative in the future to understand how exactly cleavage sites are selected, especially those downstream from classically defined promoter-proximal pausing sites. Regardless of the underlying details, our work (which built upon that of others [Skaar et al. 2015; Shah et al. 2018; Gómez-Orte et al. 2019]) has now revealed that the Integrator complex can potentially attenuate the expression of many protein-coding genes via catalyzing premature transcription termination (Elrod et al. 2019; Tatomer et al. 2019).

THE INTEGRATOR COMPLEX IS CRITICAL FOR PROPER DEVELOPMENT, PERHAPS BY CONTROLLING PROTEIN-CODING GENE EXPRESSION PATTERNS

Considering that Integrator directly controls the expression of snRNAs, enhancer RNAs, and many protein-coding genes, it is perhaps not surprising that developmental phenotypes have been observed when Integrator subunits are knocked down/out or mutated. Early work showed that mutation of IntS4 or IntS7 resulted in lethality in *Drosophila* (Rutkowski and Warren 2009; Ezzeddine et al. 2011), and a mutation in IntS6 was subsequently shown to result in embryonic defects/lethality in zebrafish (Kapp et al. 2013). Focusing specifically on the brain, depletion of Integrator subunits is associated with excess immature neuroblasts in *Drosophila* (Zhang et al. 2019) and cortical neuron migration defects in mice (van den Berg et al. 2017). Recent work has even extended these observations to humans as six individuals with severe neurodevelopmental delay have been shown to carry biallelic mutations in Integrator subunits (Oegema et al. 2017).

However, the underlying molecular mechanisms responsible for these phenotypes remain largely unclear. Slight increases in unprocessed snRNAs have been observed when Integrator subunits are depleted or mutated, but the levels of mature snRNAs have generally not been found to be significantly different (Ezzeddine et al. 2011; Oegema et al. 2017; Tatomer et al. 2019). There is no denying that functional snRNAs are critical for life, but we hypothesize that Integrator's role in fine-tuning protein-coding transcription may also be critical for normal development, and that perturbation of this tuning process may lead to the phenotypes that have been observed. For example, Integrator may catalyze premature transcription termination at genes that control differentiation processes, thereby preventing expression of the encoded proteins until they are needed. Alternatively, Integrator may be needed to repress embryonic genes in fully differentiated cells to enforce cell identity. Future studies that map Integrator target genes in different tissues as well as during differentiation programs will provide unprecedented insights into how this complex functions across the transcriptome in vivo.

SPECIALIZED INTEGRATOR SUBCOMPLEXES

Endonucleolytic cleavage by IntS11 is critical for Integrator regulation at snRNA and protein-coding genes, but our RNAi screening data indicate that these loci have different dependencies on Integrator subunits (Tatomer et al. 2019). IntS4, IntS9, and IntS11, which comprise the Integrator cleavage module (Fig. 1C; Albrecht et al. 2018), are most important for snRNA processing, whereas the noncatalytic subunits play only minor roles (Fig. 3). In contrast, large increases in mRNA expression were observed when many of the noncatalytic subunits were depleted (Fig. 3; Tatomer et al. 2019). This may suggest that distinct Integrator complexes are recruited to snRNA and

mRNA loci. Consistent with this idea, recent work has shown that IntS13 can function independently from other Integrator subunits at enhancers (Barbieri et al. 2018). Going forward, it will be critical to define molecular functions for the noncatalytic Integrator subunits, to better understand how the complex assembles, and to characterize the stoichiometry of individual Integrator subunits that are bound to target gene loci.

ATTENUATION OF EUKARYOTIC PROTEIN-CODING GENE EXPRESSION IS LIKELY MORE WIDESPREAD THAN HAS BEEN APPRECIATED

Attenuation of protein-coding gene expression via premature termination is well-established in bacteria (e.g., the tryptophan operon) (reviewed in Merino and Yanofsky 2005; Naville and Gautheret 2010), but has generally received much less attention in eukaryotes (for reviews, see Wright 1993; Kamieniarz-Gdula and Proudfoot 2019). Nevertheless, there are clear examples, some of which date back more than 40 years (Evans et al. 1979), of how regulated premature termination can modulate gene outputs in eukaryotes. In *Saccharomyces cerevisiae*, the Nrd1–Nab3–Sen1 (NNS) complex uses the Sen1 helicase to pull nascent transcripts out of the RNA polymerase II active site, thereby catalyzing premature termination of some protein-coding genes, including Nrd1 itself and genes involved in nucleotide biosynthesis (Steinmetz et al. 2001; Arigo et al. 2006; Jenks et al. 2008; Kuehner and Brow 2008; Thiebaut et al. 2008; Vasiljeva et al. 2008; Bresson et al. 2017). Notably, the NNS complex also catalyzes termination at snRNA loci, suggesting parallels with the Integrator complex (which is absent from yeast species) even though the underlying molecular mechanisms of transcription termination are distinct. In *Schizosaccharomyces pombe*, the MTREC (Mtl1–Red1 core) complex likewise triggers premature termination and promotes heterochromatin formation (Lee et al. 2013; Chalamcharla et al. 2015).

In metazoans, short capped transcripts can be detected from the 5' ends of many genes (Kapranov et al. 2007; Preker et al. 2008; Nechaev et al. 2010), likely a result of premature transcription termination events. Indeed, single-molecule footprinting (Krebs et al. 2017), ChIP-seq (Erickson et al. 2018), and fluorescence recovery after photobleaching (FRAP) experiments (Steurer et al. 2018) all suggest high levels of RNA polymerase II turnover near the 5' ends of genes. Besides Integrator, premature termination events can be driven by other complexes, such as the cleavage and polyadenylation machinery (e.g., PCF11 [Kamieniarz-Gdula et al. 2019]), the Microprocessor (Wagschal et al. 2012), and the decapping/Xrn2 machinery (Brannan et al. 2012). Nevertheless, the exact number of genes that are controlled by premature termination events is not yet clear as many prematurely terminated transcripts are rapidly degraded by the RNA exosome (Preker et al. 2008; Almada et al. 2013; Ntini et al. 2013).

WIDESPREAD PREMATURE CLEAVAGE AND POLYADENYLATION EVENTS CAN OCCUR, BUT ARE OFTEN PREVENTED BY U1 snRNP

The widespread potential for premature cleavage and polyadenylation events has been revealed in recent years via a set of analyses that began in the laboratory of Gideon Dreyfuss with a very unrelated line of scientific inquiry (for a review, see Venters et al. 2019). It was long known that stoichiometric amounts of U1, U2, U4, U5, and U6 small nuclear ribonucleoproteins (snRNPs) are required to assemble a catalytically active spliceosome (for a review, see Shi 2017), yet these snRNPs are expressed at varying levels in cells, with U1 snRNP often being most abundant. To explore the functional significance of this observation, the Dreyfuss laboratory used antisense morpholino oligonucleotides (AMOs) to functionally deplete U1 snRNA and then examined effects on the transcriptomes of human HeLa cells, mouse 3T3 cells, and *Drosophila* S2 cells (Kaida et al. 2010; Berg et al. 2012; Oh et al. 2017). Transcripts from some protein-coding genes showed accumulation of one or more introns, consistent with general splicing inhibition. However, transcripts from the majority of genes showed a strikingly distinct and unexpected pattern: the transcripts extended several kilobases from the TSS into the first intron and then abruptly ended (Fig. 6). This could have been due to U1 snRNP functioning as a

positive elongation factor (e.g., at roadblocks that cause RNA polymerase II stalling), but 3' RACE (rapid amplification of cDNA ends) revealed that these transcripts were, in fact, prematurely terminated and polyadenylated because of the use of intronic polyadenylation signals (Kaida et al. 2010; Berg et al. 2012; Oh et al. 2017). Notably, AMOs complementary to U2 snRNA or treatment with spliceostatin A, a small-molecule inhibitor of pre-mRNA splicing that binds the SF3b complex (Kaida et al. 2007), did not result in similar premature cleavage and polyadenylation patterns. This indicated that U1 snRNP has a critical role in protecting transcripts from premature cleavage that is in addition to and independent of its role in splicing (Kaida et al. 2010; Berg et al. 2012).

How does U1 snRNP control polyadenylation site usage, especially those in introns? A simple (but incomplete) model is that U1 binding to canonical 5' splice sites blocks usage of nearby cryptic polyadenylation signals. Mutation of the 5' splice site in a minigene construct inhibited splicing and resulted in increased levels of cleavage and polyadenylation in the downstream intron, but treatment with the U1 AMO increased the level of intronic termination events from this mutated construct even further (Kaida et al. 2010). Therefore, base-pairing of U1 snRNP to both canonical and noncanonical (or cryptic) 5' splice sites (Engreitz et al. 2014; Oh et al. 2017) is required for complete protection from premature cleavage

Figure 6. U1 binding to nascent transcripts controls polyadenylation signal usage. Besides canonical 5' splice sites (5'SS) at exon–intron junctions and a polyadenylation signal (PAS) at the annotated 3' end of the gene, there are often a number of cryptic 5'SS and PAS throughout nascent protein-coding transcripts. (A) Under standard conditions when U1 snRNP levels are high, U1 binds to canonical and cryptic 5'SS and blocks cleavage at nearby PAS. This results in the generation of full-length mRNAs. (B) In contrast, when U1 snRNP levels are limited, 5'SS tend to not be bound by U1, resulting in cleavage at a nearby PAS. Transcription is prematurely terminated and the generated short polyadenylated transcripts are often rapidly degraded.

and polyadenylation events (Fig. 6). At these binding sites, U1 appears to interact with a number of proteins within the cleavage and polyadenylation machinery (Boe-lens et al. 1993; Lutz et al. 1996; Gunderson et al. 1998; Awasthi and Alwine 2003), but the cleavage reaction is blocked, perhaps because of the lack of key factors that activate the cleavage reaction (e.g., CFIm68 and PABN1) (So et al. 2019).

Additional key insights into the underlying mechanism were obtained by modulating the amount of AMO that was added to cells. When U1 activity was approximately fully inhibited, cleavage and polyadenylation often occurred in the first intron with little or no transcription beyond that point. In contrast, when U1 activity was inhibited by 10%–50%, the position of the termination event was shifted further downstream from the TSS (e.g., proximal rather than distal polyadenylation sites were used in the 3' UTR) (Berg et al. 2012). Suppression of cleavage and polyadenylation by U1 is thus a 5'–3' directional process and, in general, the smaller the decrease in functional U1 levels, the greater the distance from the TSS to the polyadenylation signal that is selected for 3'-end processing. These data further suggested that 3'-end processing happens at the first actionable polyadenylation signal that is not “protected” by U1 snRNP bound nearby (within ~3.5 kb). This protective role for U1 has been referred to as “telescripting” as it is necessary for nascent transcripts to be extended over long lengths (Berg et al. 2012).

Uncontrolled premature cleavage and polyadenylation likely would result in rampant transcriptional attrition and the inability to transcribe functional mRNAs. It thus makes sense that the cell would want to only change the outputs of a subset of genes when U1 levels are limiting. In fact, it was recently shown that some genes do not undergo changes in their cleavage and polyadenylation patterns when U1 is limiting (Oh et al. 2017). Remarkably, these unaffected genes are smaller than average (median length of 14.2 kb, compared to a median length of 22.8 kb for all expressed genes) and often encode proteins with functions related to cell-stress responses or basic cellular processes required for survival, including transcription, splicing, translation, and signaling. The overall expression of some of these genes can even be up-regulated upon U1 inhibition. In contrast, genes subjected to premature cleavage and polyadenylation when U1 is limiting are longer than average (median length of 39 kb) and often encode proteins related to cell cycle progression, DNA replication, or developmental processes (Oh et al. 2017). It thus appears that evolution has broadly selected against intron expansion in genes that are crucial for cell survival under adverse conditions. This likely ensures that these genes can be more rapidly induced (as fewer nucleotides need to be transcribed) with minimal potential for premature cleavage and polyadenylation events.

CONCLUSION

Here, we have highlighted two widespread mechanisms by which premature transcription termination can be used to attenuate metazoan gene outputs—one that is catalyzed

by the Integrator complex and the other by the cleavage and polyadenylation machinery. These are near certainly not the only two mechanisms by which premature termination is catalyzed (and other mechanisms have already been identified as discussed above), but they nicely highlight how termination can shape the outputs of specific genes and thus the transcriptome. For example, U1 binding sites are depleted in the antisense direction of divergent RNAPII promoters, thus enabling cleavage and polyadenylation of antisense transcripts shortly after initiation and reinforcing promoter directionality (Almada et al. 2013; Ntini et al. 2013). Likewise, Integrator cleavage events can down-regulate the expression of some protein-coding genes by more than 100-fold. Controlling the recruitment/activity of the Integrator complex thus can likely function as a very efficient on/off gene expression switch. It is now critical for the field to reveal the underlying details of how these premature termination complexes are recruited to specific gene loci and how their cleavage activities are controlled. Potential cross talk between the premature termination pathways is unclear at this point, but seems likely to exist. In many cases, premature termination events lead to RNAs that are rapidly degraded, but it will be very interesting to determine if any of these transcripts are stable and what their biological functions are.

In total, it is becoming increasingly clear that synthesis of full-length transcripts is not necessarily the default option, and that premature cleavage events need to be actively suppressed. This raises the exciting possibility that there are many regulatory steps and fates for nascent transcripts that the field has largely not considered in the past. In addition to revealing general trends, each of these noncanonical RNAs may have its own unique biology that awaits discovery.

ACKNOWLEDGMENTS

We thank Sara Cherry, Gideon Dreyfuss, Karen Adelman, Eric Wagner, and all members of the Wilusz laboratory for discussions. This work was supported by National Institutes of Health (NIH) grants R35-GM119735 (to J.E.W.) and K99-GM131028 (to D.C.T.). J.E.W. is a Rita Allen Foundation Scholar.

REFERENCES

- Albrecht TR, Shevtsov SP, Wu Y, Mascibroda LG, Peart NJ, Huang KL, Sawyer IA, Tong L, Dundr M, Wagner EJ. 2018. Integrator subunit 4 is a ‘Symplekin-like’ scaffold that associates with INTS9/11 to form the Integrator cleavage module. *Nucleic Acids Res* **46**: 4241–4255. doi:10.1093/nar/gky100
- 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
- Arigo JT, Carroll KL, Ames JM, Corden JL. 2006. Regulation of yeast NRD1 expression by premature transcription termination. *Mol Cell* **21**: 641–651. doi:10.1016/j.molcel.2006.02.005
- Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, Hearn S, Kim Y, MacLeod AR, Krainer AR, Norton L, et al.

2016. Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes Dev* **30**: 34–51. doi:10.1101/gad.270959.115
- Awasthi S, Alwine JC. 2003. Association of polyadenylation cleavage factor I with U1 snRNP. *RNA* **9**: 1400–1409. doi:10.1261/rna.5104603
- Baillat D, Wagner EJ. 2015. Integrator: surprisingly diverse functions in gene expression. *Trends Biochem Sci* **40**: 257–264. doi:10.1016/j.tibs.2015.03.005
- Baillat D, Hakimi MA, Näär AM, Shilatifard A, Cooch N, Shiekhattar R. 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**: 265–276. doi:10.1016/j.cell.2005.08.019
- Barbieri E, Trizzino M, Welsh SA, Owens TA, Calabretta B, Carroll M, Sarma K, Gardini A. 2018. Targeted enhancer activation by a subunit of the integrator complex. *Mol Cell* **71**: 103–116 e107. doi:10.1016/j.molcel.2018.05.031
- Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**: 53–64. doi:10.1016/j.cell.2012.05.029
- Boelens WC, Jansen EJ, van Venrooij WJ, Striepecke R, Mattaj JW, Gunderson SI. 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* **72**: 881–892. doi:10.1016/0092-8674(93)90577-D
- Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, Fong N, Kiemele L, Hansen K, Davis R, Lykke-Andersen J, 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
- Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ. 2013. Dynamic integration of splicing within gene regulatory pathways. *Cell* **152**: 1252–1269. doi:10.1016/j.cell.2013.02.034
- Bresson S, Tuck A, Staneva D, Tollervey D. 2017. Nuclear RNA decay pathways aid rapid remodeling of gene expression in yeast. *Mol Cell* **65**: 787–800 e785. doi:10.1016/j.molcel.2017.01.005
- Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. 2012. Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MEN β noncoding RNAs. *Proc Natl Acad Sci* **109**: 19202–19207. doi:10.1073/pnas.1217338109
- Brown JA, Bulkley D, Wang J, Valenstein ML, Yario TA, Steitz JA. 2014. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol* **21**: 633–640. doi:10.1038/nsmb.2844
- Cazalla D, Xie M, Steitz JA. 2011. A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol Cell* **43**: 982–992. doi:10.1016/j.molcel.2011.07.025
- Chalamcharla VR, Folco HD, Dhakshnamoorthy J, Grewal SI. 2015. Conserved factor Dhp1/Rat1/Xrn2 triggers premature transcription termination and nucleates heterochromatin to promote gene silencing. *Proc Natl Acad Sci* **112**: 15548–15555. doi:10.1073/pnas.1522127112
- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33**: 717–726. doi:10.1016/j.molcel.2009.01.026
- Egloff S, Szczepaniak SA, Dienstbier M, Taylor A, Knight S, Murphy S. 2010. The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J Biol Chem* **285**: 20564–20569. doi:10.1074/jbc.M110.132530
- Elrod ND, Henriques T, Huang KL, Tatomer DC, Wilusz JE, Wagner EJ, Adelman K. 2019. The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol Cell* **76**: 738–752 e737. doi:10.1016/j.molcel.2019.10.034
- Engreitz JM, Sirokman K, McDonel P, Shishkin AA, Surka C, Russell P, Grossman SR, Chow AY, Guttman M, Lander ES. 2014. RNA–RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell* **159**: 188–199. doi:10.1016/j.cell.2014.08.018
- Erickson B, Sheridan RM, Cortazar M, Bentley DL. 2018. Dynamic turnover of paused Pol II complexes at human promoters. *Genes Dev* **32**: 1215–1225. doi:10.1101/gad.316810.118
- Evans R, Weber J, Ziff E, Darnell JE. 1979. Premature termination during adenovirus transcription. *Nature* **278**: 367–370. doi:10.1038/278367a0
- Ezzeddine N, Chen J, Waltenspiel B, Burch B, Albrecht T, Zhuo M, Warren WD, Marzluff WF, Wagner EJ. 2011. A subset of *Drosophila* integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. *Mol Cell Biol* **31**: 328–341. doi:10.1128/MCB.00943-10
- Festa RA, Thiele DJ. 2011. Copper: an essential metal in biology. *Curr Biol* **21**: R877–R883. doi:10.1016/j.cub.2011.09.040
- Gardini A, Baillat D, Cesarini M, Hu D, Marinis JM, Wagner EJ, Lazar MA, Shilatifard A, Shiekhattar R. 2014. Integrator regulates transcriptional initiation and pause release following activation. *Mol Cell* **56**: 128–139. doi:10.1016/j.molcel.2014.08.004
- Gómez-Orte E, Sáenz-Narciso B, Zheleva A, Ezcurra B, de Toro M, López R, Gastaca I, Nilsen H, Sacristán MP, Schnabel R, et al. 2019. Disruption of the *Caenorhabditis elegans* Integrator complex triggers a non-conventional transcriptional mechanism beyond snRNA genes. *PLoS Genet* **15**: e1007981. doi:10.1371/journal.pgen.1007981
- Gunderson SI, Polycarpou-Schwarz M, Mattaj JW. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell* **1**: 255–264. doi:10.1016/S1097-2765(00)80026-X
- Günther V, Lindert U, Schaffner W. 2012. The taste of heavy metals: gene regulation by MTF-1. *Biochim Biophys Acta* **1823**: 1416–1425. doi:10.1016/j.bbamcr.2012.01.005
- Hentze MW, Castello A, Schwarzl T, Preiss T. 2018. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* **19**: 327–341. doi:10.1038/nrm.2017.130
- Hernandez N. 1985. Formation of the 3' end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO J* **4**: 1827–1837. doi:10.1002/j.1460-2075.1985.tb03857.x
- Huang C, Liang D, Tatomer DC, Wilusz JE. 2018. A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs. *Genes Dev* **32**: 639–644. doi:10.1101/gad.314856.118
- Jenks MH, O'Rourke TW, Reines D. 2008. Properties of an intergenic terminator and start site switch that regulate *IMD2* transcription in yeast. *Mol Cell Biol* **28**: 3883–3893. doi:10.1128/MCB.00380-08
- Ji P, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, et al. 2003. MALAT-1, a novel noncoding RNA, and thymosin β 4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **22**: 8031–8041. doi:10.1038/sj.onc.1206928
- Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, et al. 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol* **3**: 576–583. doi:10.1038/nchembio.2007.18
- 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
- Kamieniarz-Gdula K, Proudfoot NJ. 2019. Transcriptional control by premature termination: a forgotten mechanism. *Trends Genet* **35**: 553–564. doi:10.1016/j.tig.2019.05.005
- Kamieniarz-Gdula K, Gdula MR, Panser K, Nojima T, Monks J, Wiśniewski JR, Riepsaame J, Brockdorff N, Pauli A, Proudfoot NJ. 2019. Selective roles of vertebrate PCF11 in premature and full-length transcript termination. *Mol Cell* **74**: 158–172 e159. doi:10.1016/j.molcel.2019.01.027
- Kapp LD, Abrams EW, Marlow FL, Mullins MC. 2013. The integrator complex subunit 6 (Ints6) confines the dorsal orga-

- nizer in vertebrate embryogenesis. *PLoS Genet* **9**: e1003822. doi:10.1371/journal.pgen.1003822
- Kapranov P, Cheng J, Dike S, Nix DA, Duttaputra R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, 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
- Kearse MG, Goldman DH, Choi J, Nwaezeapu C, Liang D, Green KM, Goldstrohm AC, Todd PK, Green R, Wilusz JE. 2019. Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. *Genes Dev* **33**: 871–885. doi:10.1101/gad.324715.119
- Kiledjian M. 2018. Eukaryotic RNA 5'-end NAD⁺ capping and DeNADding. *Trends Cell Biol* **28**: 454–464. doi:10.1016/j.tcb.2018.02.005
- Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE. 2015. Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev* **29**: 2168–2182. doi:10.1101/gad.270421.115
- Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. 2017. Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. *Mol Cell* **67**: 411–422 e414. doi:10.1016/j.molcel.2017.06.027
- Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. 2019. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet* **20**: 675–691. doi:10.1038/s41576-019-0158-7
- Kuehner JN, Brow DA. 2008. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol Cell* **31**: 201–211. doi:10.1016/j.molcel.2008.05.018
- Lai F, Gardini A, Zhang A, Shiekhhattar R. 2015. Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**: 399–403. doi:10.1038/nature14906
- Lee NN, Chalamcharla VR, Reyes-Turcu F, Mehta S, Zofall M, Balachandran V, Dhakshnamoorthy J, Taneja N, Yamanaka S, Zhou M, 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
- Li X, Yang L, Chen LL. 2018. The biogenesis, functions, and challenges of circular RNAs. *Mol Cell* **71**: 428–442. doi:10.1016/j.molcel.2018.06.034
- Liang D, Wilusz JE. 2014. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev* **28**: 2233–2247. doi:10.1101/gad.251926.114
- Liang D, Tatomer DC, Luo Z, Wu H, Yang L, Chen LL, Cherry S, Wilusz JE. 2017. The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. *Mol Cell* **68**: 940–954 e943. doi:10.1016/j.molcel.2017.10.034
- 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
- Lutz CS, Murthy KG, Schek N, O'Connor JP, Manley JL, Alwine JC. 1996. Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev* **10**: 325–337. doi:10.1101/gad.10.3.325
- Mao YS, Sunwoo H, Zhang B, Spector DL. 2011. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* **13**: 95–101. doi:10.1038/ncb2140
- 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
- Merino E, Yanofsky C. 2005. Transcription attenuation: a highly conserved regulatory strategy used by bacteria. *Trends Genet* **21**: 260–264. doi:10.1016/j.tig.2005.03.002
- Moore MJ, Proudfoot NJ. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**: 688–700. doi:10.1016/j.cell.2009.02.001
- Naville M, Gautheret D. 2010. Transcription attenuation in bacteria: theme and variations. *Brief Funct Genomics* **9**: 178–189. doi:10.1093/bfpg/eqq008
- Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. 2010. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**: 335–338. doi:10.1126/science.1181421
- Ntini E, Järvelin AI, Bornholdt J, Chen Y, Boyd M, Jørgensen M, Andersson R, Hoof I, Schein A, Andersen PR, 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
- Oegema R, Baillat D, Schot R, van Unen LM, Brooks A, Kia SK, Hoozeboom AJM, Xia Z, Li W, Cesaroni M, et al. 2017. Human mutations in integrator complex subunits link transcriptome integrity to brain development. *PLoS Genet* **13**: e1006809. doi:10.1371/journal.pgen.1006809
- Oh JM, Di C, Venters CC, Guo J, Arai C, So BR, Pinto AM, Zhang Z, Wan L, Younis I, et al. 2017. U1 snRNP telescripting regulates a size-function-stratified human genome. *Nat Struct Mol Biol* **24**: 993–999. doi:10.1038/nsmb.3473
- Patop IL, Wüst S, Kadener S. 2019. Past, present, and future of circRNAs. *EMBO J* **38**: e100836. doi:10.15252/embj.2018100836
- Preker R, 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
- Rubtsova MP, Vasilkova DP, Moshareva MA, Malyavko AN, Meerson MB, Zatepin TS, Naraykina YV, Beletsky AV, Ravin NV, Dontsova OA. 2019. Integrator is a key component of human telomerase RNA biogenesis. *Sci Rep* **9**: 1701. doi:10.1038/s41598-018-38297-6
- Rutkowski RJ, Warren WD. 2009. Phenotypic analysis of deflated/Ints7 function in *Drosophila* development. *Dev Dyn* **238**: 1131–1139. doi:10.1002/dvdy.21922
- Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. 2009. MENε/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci* **106**: 2525–2530. doi:10.1073/pnas.0807899106
- Shah N, Maqbool MA, Yahia Y, El Aabidine AZ, Esnault C, Forné I, Decker TM, Martin D, Schüller R, Krebs S, et al. 2018. Tyrosine-1 of RNA polymerase II CTD controls global termination of gene transcription in mammals. *Mol Cell* **69**: 48–61 e46. doi:10.1016/j.molcel.2017.12.009
- Shi Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol* **18**: 655–670. doi:10.1038/nrm.2017.86
- Shi Y, Manley JL. 2015. The end of the message: multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev* **29**: 889–897. doi:10.1101/gad.261974.115
- Skaar JR, Ferris AL, Wu X, Saraf A, Khanna KK, Florens L, Washburn MP, Hughes SH, Pagano M. 2015. The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res* **25**: 288–305. doi:10.1038/cr.2015.19
- So BR, Di C, Cai Z, Venters CC, Guo J, Oh JM, Arai C, Dreyfuss G. 2019. A complex of U1 snRNP with cleavage and polyadenylation factors controls telescripting, regulating mRNA transcription in human cells. *Mol Cell* **76**: 590–599 e594. doi:10.1016/j.molcel.2019.08.007
- Stadelmayer B, Micas G, Gamot A, Martin P, Malirat N, Koval S, Raffel R, Sobhian B, Severac D, Rialle S, et al. 2014. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat Commun* **5**: 5531. doi:10.1038/ncomms6531
- Steinmetz EJ, Conrad NK, Brow DA, Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end for-

- mation of RNA polymerase II transcripts. *Nature* **413**: 327–331. doi:10.1038/35095090
- Steurer B, Janssens RC, Geverts B, Geijer ME, Wienholz F, Theil AF, Chang J, Dealy S, Pothof J, van Cappellen WA, et al. 2018. Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA polymerase II. *Proc Natl Acad Sci* **115**: E4368–E4376. doi:10.1073/pnas.1717920115
- Sullivan KD, Steiniger M, Marzluff WF. 2009. A core complex of CPSF73, CPSF100, and Symplekin may form two different cleavage factors for processing of poly(A) and histone mRNAs. *Mol Cell* **34**: 322–332. doi:10.1016/j.molcel.2009.04.024
- Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. 2009. MEN ϵ/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19**: 347–359. doi:10.1101/gr.087775.108
- Tatomer DC, Elrod ND, Liang D, Xiao MS, Jiang JZ, Jonathan M, Huang KL, Wagner EJ, Cherry S, Wilusz JE. 2019. The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev* **33**: 1525–1538. doi:10.1101/gad.330167.119
- Thiebaut M, Colin J, Neil H, Jacquier A, Séraphin B, Lacroute F, Libri D. 2008. Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Mol Cell* **31**: 671–682. doi:10.1016/j.molcel.2008.08.010
- Tian B, Manley JL. 2017. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**: 18–30. doi:10.1038/nrm.2016.116
- van den Berg DLC, Azzarelli R, Oishi K, Martynoga B, Urbán N, Dekkers DHW, Demmers JA, Guillemot F. 2017. Nipbl interacts with Zfp609 and the integrator complex to regulate cortical neuron migration. *Neuron* **93**: 348–361. doi:10.1016/j.neuron.2016.11.047
- Vasiljeva L, Kim M, Terzi N, Soares LM, Buratowski S. 2008. Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol Cell* **29**: 313–323. doi:10.1016/j.molcel.2008.01.011
- Venters CC, Oh JM, Di C, So BR, Dreyfuss G. 2019. U1 snRNP telescripting: suppression of premature transcription termination in introns as a new layer of gene regulation. *Cold Spring Harb Perspect Biol* **11**: a032235. doi:10.1101/cshperspect.a032235
- Wagschal A, Rousset E, Basavarajaiah P, Contreras X, Harwig A, Laurent-Chabalier S, Nakamura M, Chen X, Zhang K, Meziane O, et al. 2012. Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. *Cell* **150**: 1147–1157. doi:10.1016/j.cell.2012.08.004
- Wang ZF, Whitfield ML, Ingledue TC, Dominski Z, Marzluff WF. 1996. The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev* **10**: 3028–3040. doi:10.1101/gad.10.23.3028
- Wilusz JE. 2016. Long noncoding RNAs: re-writing dogmas of RNA processing and stability. *Biochim Biophys Acta* **1859**: 128–138. doi:10.1016/j.bbtagrm.2015.06.003
- Wilusz JE. 2018. A 360 degrees view of circular RNAs: from biogenesis to functions. *Wiley Interdiscip Rev RNA* **9**: e1478. doi:10.1002/wrna.1478
- Wilusz JE, Spector DL. 2010. An unexpected ending: noncanonical 3' end processing mechanisms. *RNA* **16**: 259–266. doi:10.1261/rna.1907510
- Wilusz JE, Freier SM, Spector DL. 2008. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135**: 919–932. doi:10.1016/j.cell.2008.10.012
- 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
- Wright S. 1993. Regulation of eukaryotic gene expression by transcriptional attenuation. *Mol Biol Cell* **4**: 661–668. doi:10.1091/mbc.4.7.661
- Wu Y, Albrecht TR, Baillat D, Wagner EJ, Tong L. 2017. Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc Natl Acad Sci* **114**: 4394–4399. doi:10.1073/pnas.1616605114
- Xie M, Zhang W, Shu MD, Xu A, Lenis DA, DiMaio D, Steitz JA. 2015. The host Integrator complex acts in transcription-independent maturation of herpesvirus microRNA 3' ends. *Genes Dev* **29**: 1552–1564. doi:10.1101/gad.266973.115
- Zhang Y, Yang L, Chen LL. 2014. Life without A tail: new formats of long noncoding RNAs. *Int J Biochem Cell Biol* **54**: 338–349. doi:10.1016/j.biocel.2013.10.009
- Zhang Y, Koe CT, Tan YS, Ho J, Tan P, Yu F, Sung WK, Wang H. 2019. The integrator complex prevents dedifferentiation of intermediate neural progenitors back into neural stem cells. *Cell Rep* **27**: 987–996 e983. doi:10.1016/j.celrep.2019.03.089

Erratum

Cold Spring Harb Symp Quant Biol **84** (2019) doi: 10.1101/sqb.2019.84.039644

Erratum: Attenuation of Eukaryotic Protein-Coding Gene Expression via Premature Transcription Termination

DEIRDRE C. TATOMER AND JEREMY E. WILUSZ

The abstract of this article included an extraneous word as a result of a production proofing error upon first publication. The phrase “where which the IntS11 RNA endonuclease cleaves” should have been “where the IntS11 RNA endonuclease cleaves.” The publisher apologizes for this error. In the current version of the article, the extraneous word “which” has been removed from the abstract.

doi: 10.1101/sqb.2019.84.039784

Cold Spring Harbor Symposia on Quantitative Biology

Attenuation of Eukaryotic Protein-Coding Gene Expression via Premature Transcription Termination

Deirdre C. Tatomer and Jeremy E. Wilusz

Cold Spring Harb Symp Quant Biol published online February 21, 2020

Access the most recent version at doi:[10.1101/sqb.2019.84.039644](https://doi.org/10.1101/sqb.2019.84.039644)

Related Content

Erratum: Attenuation of Eukaryotic Protein-Coding Gene Expression via Premature Transcription Termination

Deirdre C. Tatomer and Jeremy E. Wilusz

[Cold Spring Harb Symp Quant Biol February , 2020 :](#)

P<P

Published online February 21, 2020 in advance of the print journal.

Creative Commons License

This article is distributed under the terms of the

<http://creativecommons.org/licenses/by-nc/4.0/>, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
<http://symposium.cshlp.org/subscriptions>