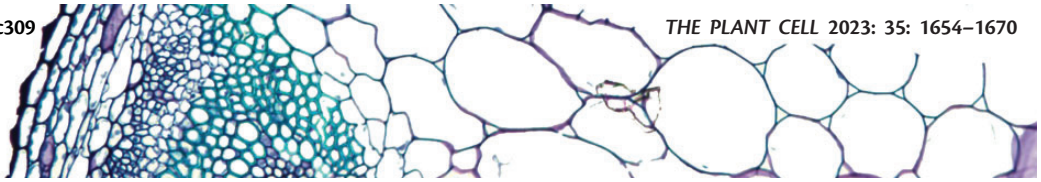
https://doi.org/10.1093/plcell/koac309 THE PLANT CELL 2023: 35: 1654–1670



Review

Cotranscriptional RNA processing and modification in plants

Sebastian Marquardt [1,*] **Ezequiel Petrillo** [2,*] **and Pablo A. Manavella** [3,*]

1 Department of Plant and Environmental Sciences, Copenhagen Plant Science Centre, University of Copenhagen, Frederiksberg, Denmark
2 Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE-CONICET-UBA), Buenos Aires, C1428EHA, Argentina
3 Instituto de Agrobiotecnología del Litoral (CONICET-UNL), Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Santa Fe 3000, Argentina

*Authors for correspondence: sebastian.marquardt@plen.ku.dk (S.M.); petry@fbmc.fcen.uba.ar (E.P.); pablomanavella@ial.santafe-conicet.gov.ar (P.A.M.)
S.M., E.P. and P.A.M. wrote and edited the manuscript and prepare the figures.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) is: Pablo A. Manavella (pablomanavella@ial.santafe-conicet.gov.ar).

Abstract

The activities of RNA polymerases shape the epigenetic landscape of genomes with profound consequences for genome integrity and gene expression. A fundamental event during the regulation of eukaryotic gene expression is the coordination between transcription and RNA processing. Most primary RNAs mature through various RNA processing and modification events to become fully functional. While pioneering results positioned RNA maturation steps after transcription ends, the coupling between the maturation of diverse RNA species and their transcription is becoming increasingly evident in plants. In this review, we discuss recent advances in our understanding of the crosstalk between RNA Polymerase II, IV, and V transcription and nascent RNA processing of both coding and noncoding RNAs.

Introduction

To perform its function, every cellular RNA molecule undergoes a maturation process. Eukaryotic RNAs undergo a series of processing steps, including capping, splicing, polyadenylation, chemical modification, slicing, re-folding, or even sequence editing, before becoming mature functional RNAs. These events not only occur to messenger RNAs (mRNAs) but also to noncoding RNA molecules and non-RNA Polymerase II (Pol II)-dependent RNAs. In metazoans, most nascent RNAs can be processed and modified while still associated with the RNA polymerase (Beyer and Osheim, 1988; Bentley, 2014; Tellier et al., 2020). The cotranscriptional nature of these processes provides a window of opportunity that allows functional coupling between them. Thus, the activity of the RNA polymerase may directly impact RNA maturation, but nascent RNAs can also, in turn, regulate transcription (Skalska et al., 2017). The coupling of transcription and RNA processing in plants is emerging as a central regulatory hub to control plant homeostasis. In the last decade, new high-throughput sequencing methods allowed the profiling of nascent transcript and cotranscriptional processing events. This review discusses recent advancements and compelling open questions in our understanding of cotranscriptional RNA processing.

RNA Pol II activity and chromatin

The DNA of eukaryotic genomes is packaged into chromatin (Luger et al., 1997). As Pol II transcription requires DNA

Received July 25, 2022. Accepted October 14, 2022. Advance access publication October 19, 2022

templates, chromatin determines key parameters of DNA processing (Gowthaman et al., 2020). Chromatin represents an obstacle to the access of Pol II to initiate transcription (Workman and Kingston, 1998). DNA packaging is particularly dense in heterochromatin, where little gene expression generally occurs. The inhibition of gene expression through altered chromatin compaction forms the basis of "epigenetic repression," where gene expression differs even though the DNA sequence is identical (Cavalli and Heard, 2019). However, heterochromatin interacts with RNAPII transcription in fascinating ways: a low level of noncoding transcription contributes to heterochromatin formation in *Schizosaccharomyces pombe* (Martienssen and Moazed, 2015), and specialized transcription complexes can direct initiation in heterochromatic DNA in fruit flies (*Drosophila melanogaster*) (Andersen et al., 2017).

Transcription factors (TFs) recognize specific DNA sequences to direct Pol II to regions in the genome where transcriptional initiation begins. Thus, Pol II transcription is inhibited when the sequences targeted by TFs are not accessible through chromatin packaging (Isbel et al., 2022). Nevertheless, following the appropriate developmental or environmental triggers, TF binding can overcome epigenetic repression mechanisms and initiate Pol II transcription (Zaret, 2020; Lai et al., 2021). Since these TFs can initiate Pol II transcription in densely packaged heterochromatin, they are also referred to as pioneer TFs. Yet precisely how pioneer TFs direct epigenome remodeling so effectively is an active area of research. In general, TFs recruit several chromatin-modifying and chromatin-remodeling activities that prepare promoter regions to initiate transcription (Haberle and Stark, 2018). Chromatin remodeling linked to basal TFs clears the promoter DNA from nucleosomes, thus creating a nucleosome-depleted region (NDR) tightly linked to the initiation of Pol II transcription (Andersson and Sandelin, 2020).

The largest subunit of Pol II (RPB1) carries a C-terminal YSPTSPS heptad repeat domain (CTD) that distinguishes Pol II from other Pol complexes (I, III, IV, and V) (Corden, 2013). Work in many organisms revealed a converging model where Pol II is recruited to promoter DNA with an unphosphorylated CTD. Data from yeast discovered paradigms that appear to apply to other systems. For example, Pol II CTD phosphorylation at the serine 5 residue (S5P) by the CTD kinase, part of the Transcription factor II H (TFIIH) basal TF complex, promotes transcription of the gene body (Komarnitsky et al., 2000). In addition, CTD S5P contributes to capping enzyme recruitment, adding the 5′-m^7G cap to nascent mRNAs (Rodriguez et al., 2000). A recent model of Pol II transcription proposes that local condensation of TFs and CTD phosphorylation contribute to phase separation to clear chromatin from gene promoters to facilitate transcriptional initiation (Cramer, 2019; Guo et al., 2019). Essentially, intrinsically disordered regions of TFs contribute to protein aggregations, offering an attractive model for the localized clustering of molecular activities to establish a cellular microenvironment where Pol II transcription through chromatin can succeed.

Pol II initiates RNA production in both directions at promoter NDRs, specifically at the edges of the +1 and −1 nucleosomes flanking of the NDRs (Andersson and Sandelin, 2020). Initiation at the −1 nucleosome edge often results in divergent long noncoding transcription (Figure 1A). The composition and patterns of posttranslational modifications of flanking nucleosomes have profound effects on the transcriptional output in each direction (Marquardt et al., 2014b): for example, through the effect of histone acetylation on the frequency of initiation (Gowthaman et al., 2021). Chromatin-based regulatory effects may explain the low extent of Pol II divergent transcription in Arabidopsis (*Arabidopsis thaliana*) compared to other eukaryotic genomes (Kindgren et al., 2020).

Interestingly, the local chromatin topology controls divergent transcription of mRNAs and their associated ncRNAs. In animals, the formation of short-range chromatin interactions, the so-called gene loops, restrict divergent transcription of bi-directional promoter NDRs enforcing transcription directionality (Tan-Wong et al., 2012). In plants, gene loop and short-range chromatin interactions influence transcriptional activity controlling the expression and identity of the mRNAs and ncRNAs produced by a genomic region (Gagliardi and Manavella, 2020). The transcriptional outcome triggered by chromatin loops largely depends on which part of a gene is encompassed within the loop. In plants, for example, this includes increased efficiency of Pol II activity (Crevillen et al., 2013; Liu et al., 2013; Gagliardi et al., 2019), transcriptional repression (Ariel et al., 2014; Kim and Sung, 2017; Guo et al., 2018; Gagliardi et al., 2019), or restricted antisense transcription (Gagliardi et al., 2019).

Initiation at the +1 nucleosome edge characterizes mRNA transcription. The +1 nucleosome provides a barrier for Pol II progression after initiation. Plant native elongation transcript sequencing (plaNET-seq) characterizes nascent RNAs bound to the RNA–DNA–RNAPII complex. This method resolves nucleosome barriers as density peaks of nascent RNAs near mRNA transcription start sites (TSSs) centered at the +1 nucleosomes (promoter-proximal stalling, Figure 1B) (Kindgren et al., 2020). Metagene analyses of posttranslational histone modification profiling data (e.g. ChIP-seq) in Arabidopsis suggest that the +1 nucleosome coincides with chromatin signatures positively correlated with mRNA expression (Leng et al., 2020b). In particular, histone acetylation and histone 3 lysine 4 tri-methylation (H3K4me3) serve as binding platforms to recruit protein complexes to promote Pol II elongation and pre-mRNA processing across species (Buratowski, 2009). The effect of chromatin-based signaling is reinforced by increasing Pol II CTD phosphorylation at the serine 2 residue (S2P) through the Arabidopsis pTEF-b kinase complex (AtCDKC2) (Kitsios et al., 2008). In several systems, CTD S2P directly recruits chromatin-modifying enzymes to facilitate elongation (Buratowski, 2009). In plants, proteomic analyses revealed

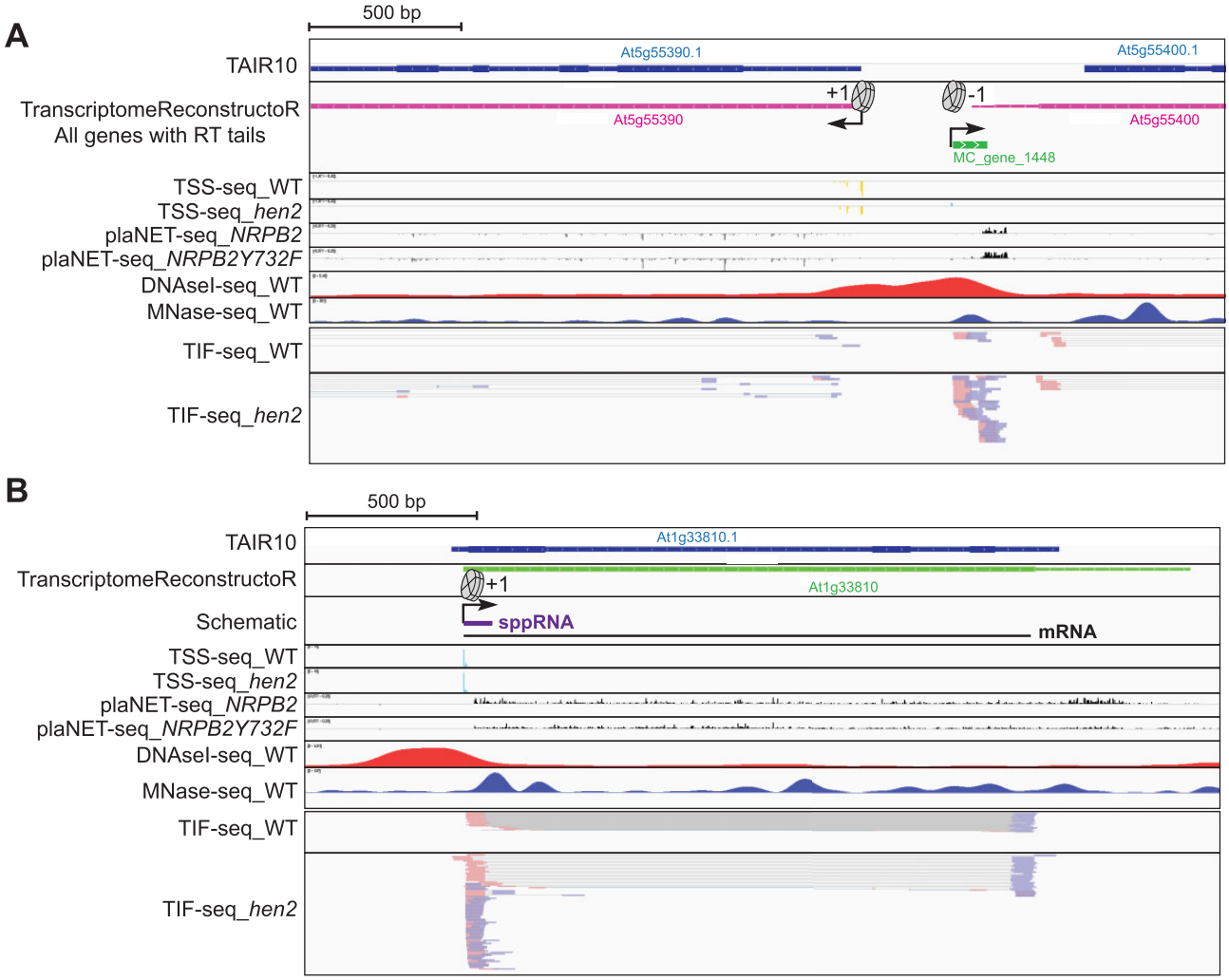


Figure 1 Arabidopsis promoter-proximal noncoding transcription. A, Divergent noncoding transcription at the At5g55390 locus. IGV genome browser screenshot of Arabidopsis transcriptomics and epigenomics data. Transcriptome annotation of TAIR10 and from the data-driven annotation tool TranscriptomeReconstructoR (Ivanov et al., 2021) are shown. TranscriptomeReconstructoR identifies MC_gene_1448 as a divergent noncoding transcript respective to At5g55390. Note that the MC_gene_1448 overlaps with the noncoding read-through tail transcript (thin pink line) of At5g55400. TSS sequencing (TSS-seq) data in Columbia-0 (Col-0) wild-type (WT) (Nielsen et al., 2019) identified the At5g55390 TSS transcribed on the crick strand (orange bar facing down). TSS-seq data in the *hen2-2* mutant (Thomas et al., 2020) identified the TSS of MC_gene_1448 (blue bar facing up). Plant native elongating transcript sequencing (plaNET-seq) data for WT NRPB2 and the NRPB2-Y732F mutant acceleration RNAII transcription are given below (Leng et al., 2020b). plaNET-seq signal for the mRNA genes encoded on the crick strand is shown (gray). plaNET-seq signal for MC_gene_1448 transcribed on the Watson strand (black) indicated strong divergent noncoding transcription. DNAseI-seq data to map accessible chromatin regions (blue) and MNase-seq to map nucleosomes (blue) are displayed below (Zhang et al., 2016). These data are consistent with the interpretation of the At5g55390 mRNA and the MC_gene_1448 divergent noncoding RNA as a divergent transcript pair from a shared promoter NDR. The positions of the flanking nucleosomes are indicated as gray cylinders. TIF-seq data for Col-0 (WT) and the *hen2-2* mutant are given below (Thomas et al., 2020). TIF-seq data reveals RNA isoforms overlapping with the MC_gene_1448 transcriptome annotation. Here, TIF-seq in the *hen2-2* mutant revealed increased levels. Divergent noncoding transcription from the promoter NDR of the At5g55390 mRNA is thus supported by data-driven transcriptome annotation, plaNET-seq tracks, TSS-seq and TIF-seq data. B, sppRNAs of mRNA at the At1g33810 locus. Data described in (A) to offer insight into transcript isoforms at the At1g33810 locus. A schematic distinguishing the sppRNA isoform (purple) from the mRNA (black) is included. mRNA and sppRNA start from the same TSS (TSS-seq, blue bar facing up), but they are distinguished by their 3′-end. Note that TIF-seq data in the *hen2-2* mutant (bottom) reveals sppRNAs as populations of RNAs ending near or shortly after initiation at the TSS, indicated by the purple TIF-seq signal (here: 3′-ends) overlapping the red signal (here 5′-ends). The sppRNA 3′-ends are positioned near the center of the +1 nucleosome indicated by the MNase-seq data (indicated by a gray cylinder). The At1g33810 locus also illustrates a separate point that reflects reduced Pol II stalling peaks in the plaNET-seq tracks of NRPB2-Y732F compared to in WT NRPB2 (black signal). Differences are most clear at the mRNA boundaries where plaNET-seq in WT NRPB2 accumulates.

the association of separate chromatin-modifying activities and elongation factors (EFs) to distinct CTD phosphorylation states (Antosz et al., 2017). Recruitment of the conserved PAF-I EF complex in several systems (Oh et al., 2004; Liu et al., 2011; Francette et al., 2021) supports the view that the pattern of CTD phosphorylation over genes assists Pol II to elongate productively through chromatin templates.

In yeast, a key function for CTD S2P during transcriptional elongation is the recruitment of histone methyltransferases that promote histone 3 lysine 36 tri-methylation (H3K36me3) (Keogh et al., 2005). This mark acts as a scaffold to recruit histone deacetylase complexes. This ensures low levels of histone acetylation in gene bodies opposite to the profile of promoter regions (Li et al., 2007). When this chromatin-based signaling system is impaired in yeast (*Saccharomyces cerevisiae*) and Arabidopsis, many cryptic intragenic TSSs are activated, resulting in additional RNA isoforms (Venkatesh et al., 2012; Nielsen et al., 2019; Wei et al., 2019; Thomas et al., 2020). These data highlight that the repressive effect of chromatin signaling positively correlates with mRNA expression on Pol II initiation within the elongation zones, perhaps indicating a mechanism to safeguard the production of full-length mRNAs.

This mechanism has important implications in genomes, for example, when an upstream Pol II transcript elongates over the promoter region of a nearby gene. In this scenario, chromatin signaling of Pol II elongation represses initiation from the downstream gene promoter (Kim et al., 2012; du Mee et al., 2018; Moretto et al., 2018; Nielsen et al., 2019). In plants, histone 3 lysine 36 di-methylation (H3K36me2), rather than H3K36me3, appears to be the chromatin signature of Pol II elongation (Leng et al., 2020a). Yet, this equivalent mechanism represses intragenic TSSs (Nielsen et al., 2019) and initiation of mRNA from promoter TSSs when these are covered by Pol II elongation from elsewhere in the genome (Nielsen et al., 2019). In conclusion, chromatin-based signaling determines the pool of cellular RNAs through effects on the initiation of competitive RNA isoforms that could be generated from the same region of DNA.

Cotranscriptional processing: the alternative splicing perspective

In eukaryotes, Pol II-derived transcripts undergo three major processing events to produce a mature RNA molecule: 5′-end capping, by adding the 5′-m^{7}G cap to nascent mRNAs; splicing, the excision of introns within the body of the precursor molecule; and the generation of a 3′-end, that usually includes polyadenylation (Proudfoot et al., 2002). Splicing and alternative splicing were discovered and reported together in two seminal papers in 1977 (Berget et al., 1977; Chow et al., 1977). Splicing is the process by which some regions of an RNA are excised (introns), while others (exons) are joined together, forming the mature RNA molecule. Splicing is directed by the splice sites (ss) at the intron/exon boundaries. While the 5′ss delimits the 5′-end of the intron, the 3′ss does the same for its 3′-end (Black, 2003).

Introns and splicing are a hallmark of eukaryotic genomes (Hawkins, 1988; Deutsch and Long, 1999). Among their possible functions, enabling alternative splicing is one of the most enticing (Gilbert, 1978). Alternative splicing is then the differential recognition of introns by the spliceosome in different RNA molecules of the same gene, giving rise to distinct alternative splicing isoforms. There are different types or categories of alternative splicing events (see Figure 2) that can be summarized as cassette exons, where an exon is included/excluded in/from the mature mRNA; alternative 5′ss, where an exon can be shortened or elongated at its 3′-end; alternative 3′ss, where an exon can be shortened or elongated at its 5′-end; and intron retention, which is different to the other possibilities as, in this case, there is no recognition and usage of different splice sites in competition, but it is a whole intron (with its 5′ and 3′ss) that is recognized and then excised or not. Since this process leads to different variants or isoforms from the same gene, it increases the transcriptome and, potentially, the proteome (Kornblihtt et al., 2013; Marquez et al., 2015; Yu et al., 2016).

The cotranscriptional nature of splicing has been beautifully shown using electron microscopy. The ultrastructural analysis of actively transcribing genes in Drosophila embryos allowed the direct visualization of nascent transcripts containing splicing loops. Moreover, some of these loops were already excised in RNA molecules located further downstream with respect to the gene (Beyer and Osheim, 1988). Different lines of evidence obtained from several organisms, including plants, reinforced this idea (Kornblihtt et al., 2013; Herzel et al., 2017; Li et al., 2020; Zhu et al., 2020; Bedi et al., 2021). In Arabidopsis, cotranscriptional splicing efficiency correlates with the level of gene expression and the distance between a given intron and the 3′-end of the gene (cleavage and polyadenylation), and it is also modulated by histone modifications, as in other organisms. Particularly in plants, it was found that cotranscriptional splicing efficiency correlates with the number of introns of a gene but is not related to the gene length (Zhu et al., 2020). Furthermore, introns with alternative 5′- or 3′-splice sites are less efficiently spliced (Li et al., 2020).

Cotranscriptionality implies that splicing and alternative splicing can be regulated by transcription. Two key models explain how Pol II transcription may regulate alternative splicing, that is, the recruitment and kinetic models, schematized in Figure 3 (Kornblihtt et al., 2013; Schor et al., 2013). The recruitment model establishes that as Pol II advances through the DNA template and synthesizes the RNA molecule, different factors are being recruited to the nascent RNA molecule and help recognize particular sequences to process introns (see Figure 3, Recruitment Model). By modulating the recruitment of specific factors during transcription, the splicing decisions might also be affected (Auboeuf et al., 2004; de la Mata et al., 2010). The kinetic model is related to Pol II speed or elongation rate. If the Pol II

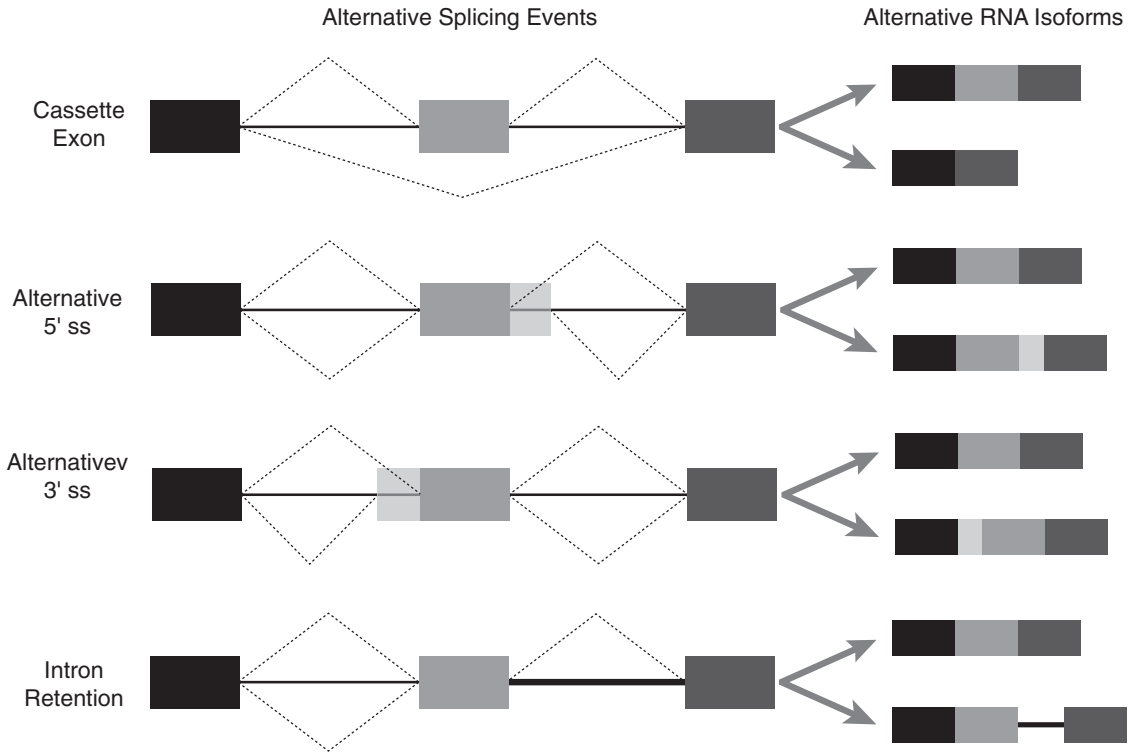


Figure 2 Main alternative splicing outcomes. Cassette exons, an exon is included/excluded in/from the mature RNA; alternative 5′ss, differential recognition and usage of distinct alternative 5′ss, mature RNAs include an exon with a 3′-end longer or shorter; alternative 3′ss, differential competitive usage of 3′ss generating isoform with a 5′-end of exons extended/shortened; intron retention is the permanence of an intron in the mature RNA. Boxes represent exons, horizontal lines specify introns, discontinuous lines indicate the recognized and used alternative splicing sites.

elongation rate is accelerated or slowed down, the splicing outcomes change. The underlying mechanism is based on the competition between different splice sites. If, for example, two splice sites are in competition and both are presented almost simultaneously, the splicing machinery will most likely choose the strongest one. However, if the transcription is slow enough, then the most proximal splice site will be synthesized by Pol II and presented to the spliceosome, earlier in time, giving the site a time window to be recognized before the other is synthesized. If the proximal site is the weakest, this time window gives it a chance to be recognized and used over the other stronger but still untranscribed splice site (Williamson et al., 2017; Dvinge, 2018).

Changes in Pol II elongation can be achieved by a net change in Pol II speed or by modulating Pol II pausing in specific regions along the DNA template. Any factor able to modulate Pol II activity is, by these means, capable of regulating alternative splicing and other processing steps that occur cotranscriptionally. Different factors regulate Pol II elongation, among them we can establish two general categories: those affecting the status of the template of transcription, the chromatin state (as explained in the previous section); and those affecting the polymerase itself, mostly the phosphorylation state of its CTD (see Figure 3, Kinetic Model). These mechanisms are deeply connected, as we discuss in the following section.

Alternative splicing regulation unveils the coupling between transcription and processing events

During the last decades, it has become increasingly evident that chromatin state and histone modifications have an impact on transcriptional elongation and, concomitantly, on alternative splicing decisions, at least for mammals (Jimeno-Gonzalez et al., 2015; Hu et al., 2017; Jabre et al., 2019). Interestingly, the coupling between transcription and splicing has also been evidenced as a dominant mechanism in plants (Henriques and Mas, 2013; Jabre et al., 2019). In recent years, the underlying mechanisms coupling Pol II elongation rate and alternative splicing outcomes in plants started to emerge.

RNA polymerases are composed of a core set of twelve subunits. NRPB2 is the conserved second-largest subunit of Arabidopsis Pol II. Interestingly, the *NRPB2Y732F* mutant renders a fast Pol II mutant that mimics its effect in yeast and shows a decreased RNA-seq intronic signal. This suggests a genome-wide trend toward increased splicing efficiency in plants when Pol II transcription is accelerated. This is in opposition to reports in yeast and Drosophila, where splicing efficiency is increased with a slow elongation rate (Khodor et al., 2011; Braberg et al., 2013; Aslanzadeh et al., 2018). Furthermore, in mammalian cells, a higher or a lower elongation rate, compared to the wild-type condition, causes higher levels of intron retention and an enhanced inclusion or skipping of specific alternative exons (Fong et al., 2014; Maslon et al., 2019). Hence, perturbing proper Pol II

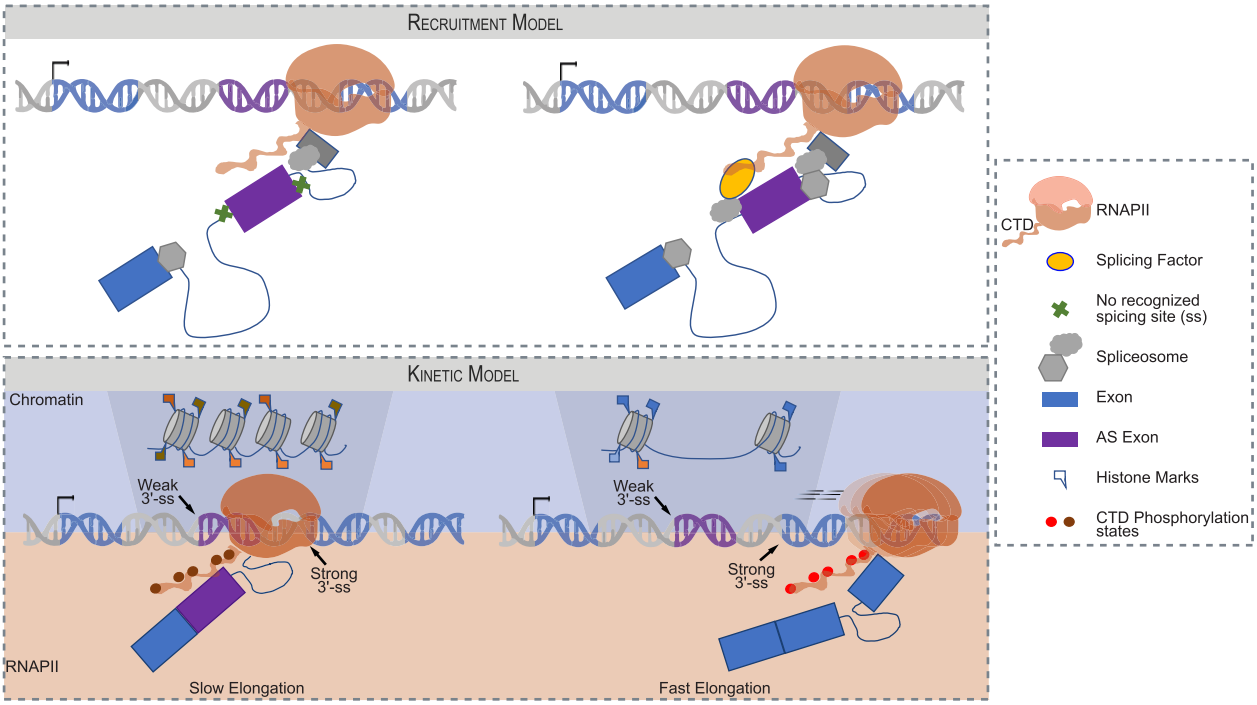


Figure 3 Impact of RNA polymerase II transcription on alternative splicing decisions. The Recruitment Model (top) is based on the recruitment of splicing factors mediated by interactions with the Pol II CTD. Binding of SR proteins could favor the usage of different splice sites in competition, modulating alternative splicing outcomes. In the Kinetic Model (bottom), the regulation of alternative splicing is modulated by the rate of Pol II elongation (de la Mata and Kornblihtt, 2006). Slow elongation (left) gives more time for recognition of weak splice sites, while fast elongation results in the usage of the stronger sites as a result of the competition (right). The rate of elongation can be influenced, among other things, by the chromatin state that can present obstacles to Pol II movement and by the level of phosphorylation of different residues in the repeats of the heptapeptide unit Y1S2P3T4S5P6S7 (42 in Arabidopsis; Dietrich et al., 1990) of the large Pol II-subunit CTD. Boxes represent exons, lines represent introns.

elongation has dramatic and specific consequences in RNA processing. In this sense, the *NRPB2Y732F* mutant plants also show higher levels of intron retentions and changes in the inclusion of particular cassette exons, with a tendency toward an increase in the usage of proximal 5′ss and distal 3′ss (Leng et al., 2020b). This was consistently reported, again in Arabidopsis, using a mutant of *NINETEEN COMPLEX-RELATED PROTEIN 1*, which encodes an accessory spliceosomal component that facilitates spliceosome intron lariat release, showing a Pol II occupancy at genes consistent with a fast Pol II elongation, similar to the *NRPB2Y732F* mutant.

Building upon this, when Pol II pauses more often, the effect on alternative splicing goes, in general, in the expected opposite direction if considering a kinetic regulation of this process. This is the case when expressing an enhanced mutant for the elongation factor TFIIS, with two key amino acids of the trigger loop changed (Sigurdsson et al., 2010). This TFIISmut renders a protein that blocks Pol II endonucleolytic cleavage in vitro, causing it to pause more often (Sigurdsson et al., 2010). This mutated factor changes the alternative splicing patterns of several genes in Arabidopsis, with a strong tendency toward enhanced intron splicing and an increase in exon inclusion (Dolata et al., 2015). These results are similar to the observations in animal cells using a slow Pol II mutant, the C4 point mutation in its largest subunit (de la Mata et al., 2003). Even though average intron length and some features of introns and exons substantially differ between animals and plants, these findings allow us to conclude that the coupling of splicing and transcription is conserved in plants as well as the key factors of the splicing machinery.

Light, one of the most vital environmental determinants for plant development, promotes transcriptional elongation in a set of genes whose alternative splicing events are also regulated by this environmental cue (Godoy Herz et al., 2019). Moreover, these effects are lost when using the TFIISmut transgenic line, indicating that both processes are mechanistically coupled (Godoy Herz et al., 2019). Interestingly, the alternative splicing response to light/dark transition is altered in chromatin remodeling mutants (i.e. *hd1* mutant for histone deacetylation) or pharmacological treatments that affect chromatin compaction (i.e. TSA, an inhibitor of histone deacetylases) (Godoy Herz et al., 2019). This reinforces the notion that the chromatin state controls alternative splicing decisions by modulating the Pol II elongation rate.

Temperature is another environmental cue that profoundly affects plant development. In animals, Pol II activity at the beginning and end of genes is altered to modulate

the transcriptional output to cope with temperature changes (Rougvie and Lis, 1988; Bunch et al., 2014). Analyses of temporal dynamics of nascent Pol II transcription in response to cold using a NET-seq approach in Arabidopsis showed changes in Pol II promoter-proximal stalling and at the 3′-end (Kindgren et al., 2020). In addition, the data suggest that cold modulates the Pol II elongation rate in gene bodies and that splicing represents a transient transcriptional barrier at low temperatures in plants. Interestingly, Pol II tends to stall 25 nucleotides into introns, and the amplitude of this peak correlates with intron length (Kindgren et al., 2020). Hence, the temperature is another signal that finely regulates Pol II elongation to adjust gene expression to keep up with changes in environmental conditions.

In plants, intron retention is reported as the most frequent alternative splicing event (Ner-Gaon et al., 2004; Marquez et al., 2012; Zhang et al., 2022). However, determining whether this is an effective alternative splicing outcome is not trivial. This is mainly due to the common assumption that capped and poly(A)-containing mRNAs that retain one out of several introns are "fully"-processed mature molecules, when there is no certainty if the retained intron can be/is spliced afterward (posttranscriptionally). It was recently shown in Arabidopsis, using nanopore sequencing technology, that about half of the introns remain unspliced after Pol II synthesizes 1 kb downstream of the 3′ss. Moreover, many full-length chromatin-bound RNA molecules are polyadenylated yet still contain unspliced introns at specific positions. These results suggest that intron retention is overestimated in plants and that posttranscriptional excision of these introns could generate the reference isoforms when needed or in response to particular conditions. Such a mechanism would be a widespread form of intron detention, which is a specific event of intron retention where splicing resumes upon a signal or cellular activation (Boothby et al., 2013; Boutz et al., 2015).

Since intron retention events attracted most attention in plants, the other options have likely been functionally underrated (Martin et al., 2021). Long-read-sequencing technologies promise to increase our knowledge about other events and their functions in different cell types and conditions and, more importantly, in other plant species.

Termination and 3′-end formation

Termination of Pol II transcription offers a key mechanism to diversify the populations of RNA molecules from a single gene. Alternative polyadenylation (APA) represents a classic cellular mechanism to inhibit the expression of full-length mRNA isoforms through the selection of a polyadenylation (p(A)) site earlier in the gene (Kamieniarz-Gdula and Proudfoot, 2019). In plants, APA is a well-defined cotranscriptional RNA processing event (Yang et al., 2021) that controls many aspects of the plants' life, including development, stress responses, hypoxia, and seed dormancy (Cyrek et al., 2016; de Lorenzo et al., 2017; Yu et al., 2022). During the regulation of flowering time, the production of full-length mRNA of several genes encoding RRM domain RNA-binding proteins, such as FLOWERING CONTROL LOCUS A (FCA) and FLOWERING CONTROL PROTEIN A (FPA), are balanced by APA in large introns (Quesada et al., 2003; Hornyik et al., 2010; Parker et al., 2021). Interestingly, the FCA/FPA proteins encoded by the full-length mRNA isoforms stimulate intronic APA, highlighting a negative feedback loop through APA.

While genetic data supported the hypothesis that FCA/FPA affects flowering through regulation of the central flowering repressor gene *FLOWERING LOCUS C* (*FLC*), molecular support for APA at *FLC* remained elusive until the *FLC* antisense lncRNA *COOLAIR* was identified (Liu et al., 2010). Reminiscent of alternative mRNA processing, *COOLAIR* expresses several lncRNA isoforms through APA and alternative splicing (Liu et al., 2010; Marquardt et al., 2014a; Li et al., 2015). Characterizing the molecular mechanism of FCA-mediated *FLC* repression revealed a correlation with APA of *COOLAIR*, rather than the *FLC* mRNA. Subsequent genetic and molecular dissection elaborated on the chain of events connecting APA of the lncRNA *COOLAIR* to chromatin-based repression of *FLC* transcription (Wu et al., 2020; Zhao et al., 2021).

When APA occurs in an intron, efficient splicing could eliminate the sequences targeted by APA, thus setting up a competition between APA and splicing. The Nanopore Direct RNA sequencing technology allows powerful insight into genome-wide APA in Arabidopsis (Zhang et al., 2020; Parker et al., 2021). Molecular characterization of suppressor mutations and natural allelic variations at *FLC* suggest that APA at *COOLAIR* is in functional competition with splicing to affect flowering time (Marquardt et al., 2014a; Li et al., 2015). A natural accession of Arabidopsis with a causative genetic variation at *FLC* displayed noncoding polymorphisms that define *COOLAIR* splicing isoforms (Li et al., 2015). These data indicate that genetic polymorphisms facilitating adaptation to local environments affect RNA processing of a lncRNA, which functionally impacts mRNA expression (Zhao et al., 2021). This is consistent with functional studies of *COOLAIR* in the context of natural temperature fluctuations that highlighted a key role of *COOLAIR* transcripts in regulating *FLC* expression in natural temperature fluctuations (Zhao et al., 2021).

It is important to highlight that pre-mRNA cleavage at the 3′-ends of eukaryotic Pol II mRNA transcripts results in two RNA molecules: (1) the pre-mRNA that carries a 5′-m^7G cap and a 3′-end that will receive the p(A) tail (i.e. mRNA) and (2) a noncoding RNA isoform that lacks a 5′-cap and that is still elongated by Pol II (i.e. read-through tail) (Nojima and Proudfoot, 2022). Work in mammalian cells and yeast suggested the torpedo mechanism for transcriptional termination; it posits that Pol II transcription continues until a 5′- to- 3′-XRN exonuclease catches it and dislodges the polymerase from the template (Kim et al., 2004; West et al., 2004; Proudfoot, 2016). Of note, most commonly used transcriptomic methods rely on the 5′-m^7G cap or p(A) purifications, thus read-through lncRNA

isoforms found at all mRNA genes are missing in genome annotations unless high-quality data-driven annotation tools are used (Ivanov et al., 2021) (Figure 1).

Reminiscent of the competition between splicing and APA, a kinetic competition between post-p(A) site, Pol II transcription, and the XRN 5′-exonuclease defines the efficiency of transcriptional termination in yeast (Hazelbaker et al., 2013). Point mutations in Pol II increasing transcription speed often result in extended read-through transcription also in Arabidopsis (Leng et al., 2020b) (Figure 1B). This observation suggests that accelerated Pol II can outrun the XRN exonuclease delaying transcription termination.

The lncRNA isoforms from read-through transcription can affect the initiation of neighboring mRNAs through chromatin-based repression linked to Pol II elongation. Specifically, the read-through isoform (*asCBF1*) of the lncRNA *SVALKA* promotes Arabidopsis cold resilience by limiting the duration of *C-REPEAT/DRE BINDING FACTOR 1* (*CBF1*) expression (Kindgren et al., 2018). Over-expression of *CBF* genes increases the resilience of cold-sensitive crops to cold, yet plant growth is limited, resulting in yield penalties (Gilmour et al., 2000; An et al., 2016). Feedback repression of *CBF1* by the *SVALKA-asCBF1* cascade may illustrate how plant genomes limit the *CBF* expression with lncRNA-based feedback repression to short intervals, thus harnessing the benefits of increased cold tolerance without associated fitness penalties. In conclusion, even though the lncRNA isoforms resulting from RNA 3′-end formation are missing in most genome annotations, they can be important to assist plants in plant–environment interactions.

Metagene representation of plaNET-seq data highlighted a large density of peaks at the 5′- and 3′- boundaries of genes encoding mRNAs (Kindgren et al., 2020). These peaks, which indicate stalled Pol II complexes, support the view that nascent Pol II transcripts accumulate before elongation progress and termination occur. Consistently, plaNET-seq experiments assaying accelerated Pol II transcription revealed a strong reduction of Pol II-peaks at gene boundaries, supporting the view that a transcription slow-down at gene borders contributes to these accumulations (Leng et al., 2020b). It is intuitive to imagine how Pol II stalling at 3′-ends facilitates transcriptional termination, for example, by facilitating XRN 5′- to 3′-exonuclease to catch up with Pol II complexes engaging in read-through transcription. Conversely, the role of Pol II low transcription at 5′-ends is less clear. In Arabidopsis, the position of promoter-proximal Pol II stalling maps to the center of the +1 nucleosomes, reinforcing the view of chromatin-based signal integration here (Kindgren et al., 2020).

An attractive model inspired by metazoan data is to interpret promoter-proximal Pol II stalling as assembled and elongation-competent Pol II complexes that are stored on genes in a poised state (Core and Adelman, 2019). Poised Pol II complex could provide a mechanism to accelerate environmentally triggered transcriptional adjustments since the steps resulting in transcriptional initiation are already completed (Adelman et al., 2005). While this model offers an intriguing rationale for promoter-proximal Pol II stalling, a time course of cold-triggered plant gene induction using plaNET-seq that could resolve poised complexes offered little evidence for mRNA induction based on the release of poised Pol II complexes from +1 nucleosome reservoirs in these conditions (Kindgren et al., 2020). It will be interesting to resolve if other environmental responses of plants that rely on rapid transcriptional gene induction offer more clear support for plant mRNA induction by releasing poised Pol II complexes.

An alternative model for the role of promoter-proximal Pol II stalling in Arabidopsis posits that these peak densities indicate promoter-proximal termination of Pol II transcription (Thomas et al., 2020). Transcript isoform sequencing (TIF-seq) is a method based on short-read sequencing to resolve the 5′- and 3′-ends of individual RNA molecules (Pelechano et al., 2013). TIF-seq data identified a class of noncoding RNA that shares the TSSs with mRNA isoforms but terminates around 100 nucleotides after the TSSs, named short promoter-proximal RNAs (sppRNAs) (Thomas et al., 2020) (Figure 1B). The 3′-end position of sppRNAs coincides with the peak of promoter-proximal stalling at +1 nucleosomes. In addition, sppRNA-producing Arabidopsis genes have a higher promoter-proximal Pol II stalling signal. These data suggest that promoter-proximal Pol II stalling is linked to RNA 3′-end formation shortly after initiation, resulting in sppRNAs. At first glance, this situation may resemble a mechanism to repress full-length mRNA production similar to APA, where transcriptional induction yields sppRNAs instead of full-length mRNAs. However, sppRNA formation positively correlates with mRNA induction, arguing against a repressive role of sppRNAs regarding full-length mRNA isoform production. Moreover, the elimination of DNA sequences correlating with sppRNA formation reduces gene expression in heterologous reporter assays, perhaps indicating a stimulating effect of sppRNA production on gene induction during coupling of both processes. While the roles of sppRNA formation in mRNA induction remain to be clearly defined, it is conceivable that short noncoding RNA isoforms could contribute to the condensation of chromatin remodeling activities and TFs to facilitate gene induction (Sharp et al., 2022). In conclusion, the discovery of sppRNAs offers a model to reconcile the two characteristic Pol II stalling sites at plant genes with functions in 3′-end formation.

Cotranscriptional and posttranscriptional effects of RNA modifications

Recently, RNA base modifications have attracted attention in different systems, including plants. This process can be compared to DNA modifications, as the epitranscriptome can also be dynamically written, erased, and read, and each of these events can modulate gene expression. As is the case for the other mechanisms revisited here, our knowledge of plant RNA modifications is far more limited than that of

animals. It is known that N^6-methyladenosine (m^6A) is the most prevalent internal RNA modification in eukaryotes, bacteria, and viruses. Plants are not an exception. From more than 100 known RNA modifications, besides m^6A, only uridylation and m^5C have been identified in plants so far (Xiao et al., 2016; David et al., 2017; De Almeida et al., 2018; Kumar and Mohapatra, 2021). RNA modifications on stable mRNAs different from m^6A were found mostly associated with alternatively spliced introns (Vandivier et al., 2015), once again suggesting regulation of alternative splicing decisions directed by these modifications. On the other hand, m^5C modifications in Arabidopsis accumulate in the CDS and are particularly enriched in mobile mRNAs (Yang et al., 2019; Zhang et al., 2020).

Within the spirit of this review, the key question is whether these modifications occur cotranscriptionally, and if so, whether this process is coupled with transcription. In this sense, it is relevant to mention that m^6A can modulate splicing through its reader, YTH Domain-Containing 1 (YTHDC1) (Xiao et al., 2016). Briefly, YTHDC1 interacts with several pre-mRNA splicing factors, including five serine–arginine-rich proteins. Interestingly, while YTHDC1 and SRSF3 (SRp20) mainly contribute to exon inclusion, SRSF10 primarily induces exon skipping (Xiao et al., 2016). This report indicates that RNA modifications can occur cotranscriptionally affecting downstream processing events. In this sense, m^6A modifications deposited cotranscriptionally near splice junctions affect splicing kinetics. Intronic internal m^6A signal is associated with slow processing and alternative splicing in human cells in culture. More specifically, intronic m^6A peaks are associated with upstream or downstream exon skipping (Louloupi et al., 2018). An interesting example of m^6A deposition in a cotranscriptional manner comes from *Caenorhabditis elegans*. The deposition of this mark on the 3′ss of the S-adenosylmethionine synthetase pre-mRNA inhibits its proper splicing and protein production in response to a rich diet (Mendel et al., 2021). On the other hand, m^6A in Arabidopsis is enriched not only around the stop codon and within 3′-untranslated regions (3′-UTRs), but also around the start codon.

Gene ontology analysis indicates that this unique distribution is associated with plant-specific pathways involving the chloroplast (Luo et al., 2014). Loss of m^6A from 3′-UTRs is associated with decreased relative transcript abundance and defective 3′-end formation (Parker et al., 2020). This is consistent with evidence suggesting that m^6A predominantly protects Arabidopsis mRNAs from endonucleolytic cleavage (Anderson et al., 2018), which is opposed to the effect in human cells where this modification promotes decay (Wang et al., 2014). Remarkably, plant YTH domain proteins that bind m^6A-modified RNAs, mainly EVOLUTIONARILY CONSERVED C-TERMINAL REGION 2 and 3 (ECT2 and ECT3), do not seem to affect polyadenylation site choice in their target mRNAs as previously thought. Though their inactivation leads to lower steady-state accumulation of their target transcripts, this is most likely due to cytoplasmic effects (Arribas-Hernandez et al., 2021a, 2021b). Hence, this mechanism is not contributing to cotranscriptional 3′-end processing decisions.

RNA modifications have cotranscriptional consequences, such as changing alternative splicing patterns, and posttranscriptional effects, such as regulating stability and mRNA mobility, which are also connected to other processing events. Of relevance, the disruption of the m^6A writer, the only mRNA adenosine methylase (MTA) in Arabidopsis, leads to an arrest in embryo development. Interestingly, this MTA-deficient line has decreased levels of microRNAs (miRNAs) but accumulates primary miRNA transcripts (pri-miRNAs). Moreover, pri-miRNAs are methylated by MTA, and RNA structure probing analysis reveals a decrease in secondary structure within stem–loop regions of these transcripts in the MTA-deficient plants (Bhat et al., 2020). Interestingly, this report showed that MTA interacts with Pol II through TOUGH (TGH). TGH is also required to efficiently recruit the core miRNA biogenesis complex to the transcriptional machinery. These findings suggest that m^6A methylation of pri-miRNAs is likely cotranscriptional, linking the small RNA pathways with transcription.

Small RNA production at the chromatin level

The production of transposable element (TE)-derived sRNAs (the so-called heterochromatic small interfering RNAs [siRNAs], hetsiRNA) is long-known to be coupled to transcription. This process, which ultimately leads to de novo DNA methylation through the RNA-directed DNA methylation (RdDM) pathway, initiates with the transcription, by the plant-exclusive Pol IV, of TE and repetitive elements. In a transcriptionally coupled event, the short nascent RNAs are converted into dsRNAs by the RNA-dependent RNA Polymerase 2 (RDR2) in Arabidopsis (Haag et al., 2012; Singh et al., 2019; Mishra et al., 2021). The direct interaction of Pol IV with the RDR2 positions this enzyme to rapidly engage the free 3′-ends of Pol IV transcripts, converting them into double-stranded RNAs (Mishra et al., 2021). Interestingly, the Pol IV is not only necessary to recruit RDR2, but its transcriptional activity directly influences the RdDM pathway. In this sense, while the CTD of Pol IV's largest subunit is dispensable for the complex catalytic activity and Pol IV termination-dependent activation of RDR2, the levels of 24-nt siRNA decrease ∼80% when the CTD is deleted (Wendte et al., 2019). This observation strengthens the importance of the coupling between transcription and the RdDM pathway.

After the relatively short nascent RNAs are converted into dsRNA by RDR2, DICER-LIKE 3 (DCL3) recognizes them to produce 23- to 24-nt het-RNAs (Loffer et al., 2022). Whether the association of DCL3 with the Pol IV-dependent dsRNAs also occurs coupled with transcription remains to be addressed. On one side, we can speculate that this is a feasible scenario considering that DCL3 shares protein partners with DCL1, such as DAWDLE and PLEIOTROPIC REGULATORY LOCUS 1, which are likely to associate with the chromatin and act cotranscriptionally (Zhang et al., 2014, 2018; Gonzalo et al., 2022; Mencia et al., 2022). On the

other hand, the Pol IV transcripts are relatively short, ~30 nt, which implies that it is probable that these transcripts are rapidly transcribed, converted to dsRNA, and released before any further processing can occur (Singh et al., 2019).

Independently of where DCL3 acts, the outcome of its processing is a pool of 23- to 24-nt siRNAs that, upon loading into ARGONAUTE 4 or 6 (AGO4 or AGO6), fulfill their functions guiding the DNA methylation machinery to specific loci (Duan et al., 2015; Wang and Axtell, 2017). The methylation of a target locus, normally the same that originated the siRNAs in the first place, is also coupled to transcription. Here, AGO4/6 recognizes and slice Pol V-nascent transcripts that act as scaffolds to recruit DNA methyltransferases to the locus triggering its de novo methylation (Liu et al., 2018). This coupling between Pol V transcription and AGO4 recognition of the nascent RNA is a *sine-qua-non* requisite for RdDM as it allows the positioning of AGO4 and the consequent recruitment of the methylation machinery adjacent to the target locus. However, recent evidence indicates that AGO4 targeting precedes transcription and is necessary for the initial recruitment of Pol V, which will then trigger the cycle of RdDM at a transcribed target locus in Arabidopsis (Sigman et al., 2021).

Once initiated, RdDM is a feed-forward cycle that reinforces DNA methylation at target loci as both Pol IV and Pol V are recruited to methylated loci. However, how this process initiates at nonmethylated loci was a long-standing open question that we only recently started understanding (Hung and Slotkin, 2021). Based on current evidence, TE-derived Pol II transcripts can trigger the production of siRNAs to initiate RdDM (Sigman et al., 2021). These siRNAs can, upon loading in AGO4, recognize target loci, prime Pol V's recruitment, and trigger the first round of methylation that will ultimately promote Pol IV association to fuel the RdDM feed-forward loop (Sigman et al., 2021) (Figure 4). However, a recent report suggests that Pol V produces low levels of nonspecific transcripts that can initiate RdDM when matching available 24-nt siRNAs produced either in trans or in cis (Tsuzuki et al., 2020) (Figure 4). It is also likely that the first round of methylation is not triggered by 24-nt hetsiRNAs, at least in cis, but by 21–22-nt epigenetically active siRNAs (easiRNAs) produced after the miRNA cleavage of TE transcripts (Creasey et al., 2014; Ariel and Manavella, 2021).

Besides the well-studied coupling between Pol IV and Pol V activities and the canonical RdDM pathway, the methylation of DNA triggered by siRNAs can also depend exclusively on Pol II transcripts. This is the case of siRNA produced from highly complementary stem–loop structured transcripts and other Pol II-derived TE transcripts found at least in Arabidopsis and sunflower (*Helianthus annus*) (Sasaki et al., 2014; Panda et al., 2016; Gagliardi et al., 2019) and for DNA-methylation initiated by DCL-independent siRNAs in Arabidopsis (sidRNAs; Ye et al., 2016). In the first case, coupling between transcription and siRNA production is likely to occur as the production of these siRNAs resembles miRNA biogenesis, a process that, as discussed below, occurs cotranscriptionally. In the case of sidRNAs, although most of these molecules are produced in a Pol IV/RDR2 mechanism reminiscent of the canonical RdDM pathway, a fraction of them relies on Pol II transcripts.

The proposed model indicates that precursor transcripts produced by the Pol II from TE loci can associate with AGO4. These transcripts are then subject to 3′- to 5′-exonucleolytic trimming for maturation, reaching the canonical 24-nt size. Then, these sidRNAs can trigger de novo methylation of the locus priming the recruitment of Pol IV to amplify the locus silencing by the canonical RdDM pathway (Ye et al., 2016). This offers interesting parallels to *S. pombe*, where RNAi relies on Pol II transcripts to form heterochromatin. While RNAi directed to mRNA genes is inefficient, mutations in the PAF-I complex enhance its efficiency (Kowalik et al., 2015). The data are consistent with the idea that PAF-I recruitment, which is linked to phosphorylation at the Pol II CTD, prevents heterochromatin formation, perhaps by favoring productive elongation. It is tempting to speculate that the loss of the NRPB1-like CTD in Pol IV and Pol V (Haag and Pikaard, 2011), and thus reduced capacity to recruit Pol II-like EFs, could have been key drivers in the specialization of Pol IV and Pol V to participate in transcription for slicing.

Coupling miRNA processing with Pol II transcription

Recent evidence in Arabidopsis suggests that miRNA processing is also coupled to transcription in plants. The earliest evidence of a potential coupling between transcription and miRNA processing came from the discoveries that several proteins known to associate and modulate the Pol II activity, such as C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1, the MEDIATOR complex, NEGATIVE ON TATA LESS2, or PROTEIN PHOSPHATASE 4, among others, could also regulate miRNA biogenesis in plants (Kim et al., 2011; Manavella et al., 2012; Wang et al., 2013, 2019). Similarly, many proteins regulating splicing, a well-known cotranscriptional event, also affect miRNA production, for example, MOS4-ASSOCIATED COMPLEX 7 (MAC7), SERRATE (SE), SERINE/ARGININE-RICH SPLICING FACTOR 40/41, GLYCINE-RICH RNA-BINDING PROTEIN 7, or the CAP-Binding Complex (Laubinger et al., 2008; Koster et al., 2014; Raczynska et al., 2014; Chen et al., 2015; Jia et al., 2017). In fact, intron distribution and length directly affect the efficiency of miRNA precursor processing, suggesting a coupling between miRNA biogenesis and splicing (Bielewicz et al., 2013; Szweykowska-Kulinska et al., 2013; Knop et al., 2017).

Even when it was tempting to speculate that miRNA biogenesis was coupled to Pol II activity, it was plausible that the identified proteins have dual and independent functions during transcription and miRNA biogenesis. The discovery in 2015 that the Elongator complex, which interacts with Pol II, allows the recruitment of DCL1, the RNase responsible for miRNA production, directly to the *MIRNA* encoding loci

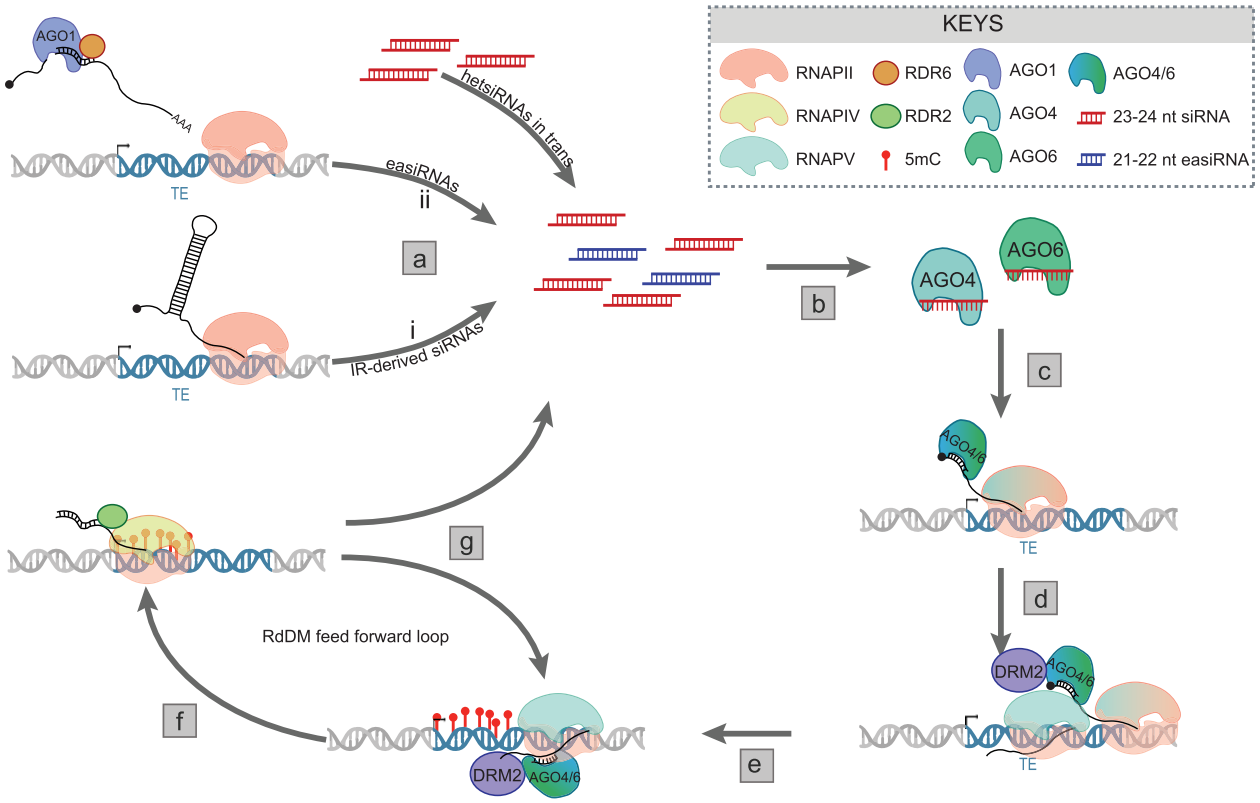


Figure 4 Proposed model of RdDM initiation. Unmethylated loci containing TEs can be transcribed by Pol II to produce siRNAs from highly complementary dsRNA hairpin transcripts derived from inverted repeated sequences (Sasaki et al., 2014) (a.i) or easiRNAs after the initial cut of the TE transcripts by an miRNA-loaded AGO1 and conversion to dsRNA by RDR6 (Creasey et al., 2014) (a.ii). A nuclear pool of hetsiRNA is then preferentially loaded into AGO4 or AGO6 (b). AGO-loaded siRNAs will recognize unmethylated target loci either by binding Pol II transcripts (Sigman et al., 2021), scarce transcripts produced by surveilling Pol V (Tsuzuki et al., 2020), or even perhaps by direct AGO–DNA interaction (Lahmy et al., 2016) (c). AGO4/6 then act as a scaffold recruiting Pol V (Sigman et al., 2021) (d), which in turn initiates de novo methylation of the targeted loci (e). Methylated DNA will promote the recruitment of Pol IV to the locus (f) triggering the production of RDR2/DCL3-derived siRNAs and initiating the feed-forward loops of RdDM that will keep the locus silenced (g).

provided direct evidence of the microprocessor association with chromatin (Fang et al., 2015). After this initial discovery, several other components of the miRNA processing machinery, such as HASTY (HST), HYPONASTIC LEAVES 1 (HYL1), TGH, and MAC7, were found to be associated with the *MIRNA* loci (Jia et al., 2017; Bhat et al., 2020; Tomassi et al., 2020; Cambiagno et al., 2021). These reports strengthen the idea of cotranscriptional processing of miRNAs in plants.

Still, the recruitment of the processing machinery to *MIRNA* loci does not necessarily imply a coupling between transcription and the processing of the miRNA precursors. It could be argued that given the extensive length and complex secondary structures of plant miRNA precursors, even when the microprocessor is recruited to the loci, it is likely that transcription ends before the processing could start. Unlike mammals, where direct evidence showed that the first processing step by Drosha occurs cotranscriptionally (Morlando et al., 2008; Nojima et al., 2015), we lacked such direct evidence in plants. Only in 2022, by sequencing nascent transcripts, we detected miRNA-processing fragments still associated with Pol II, demonstrating the coupling between transcription and miRNA biogenesis (Gonzalo et al., 2022). Cotranscriptional processing of miRNA precursors processed from the loop occurs entirely coupled to transcription. In contrast, those pri-miRNAs processed from the base undergo a second nucleoplasmic dicing step. This difference suggests that the processing complex and the transcriptional machinery remain bound through their interaction with the nascent pri-miRNA (Figure 5).

Still, one of the most outstanding open questions is how the *MIRNA* loci are specifically recognized from any other gene to recruit the processing machinery during transcription. In this sense, whether a miRNA undergoes cotranscriptional processing appears to be dictated by R-loops within the *MIRNA* loci (Gonzalo et al., 2022). R-loops are by themselves cotranscriptional features as these RNA/DNA hybrid structures are well-known to form during transcription (Chedin, 2016; Kim and Wang, 2021). Specific proteins likely recognize these structures facilitating the recruitment of the processing complex to the *MIRNA* loci. This could be the case of FCA at the *FLC* locus, where the protein association resolves cotranscriptional R-loops triggering chromatin silencing in Arabidopsis (Xu et al., 2021). Thus, R-loops may

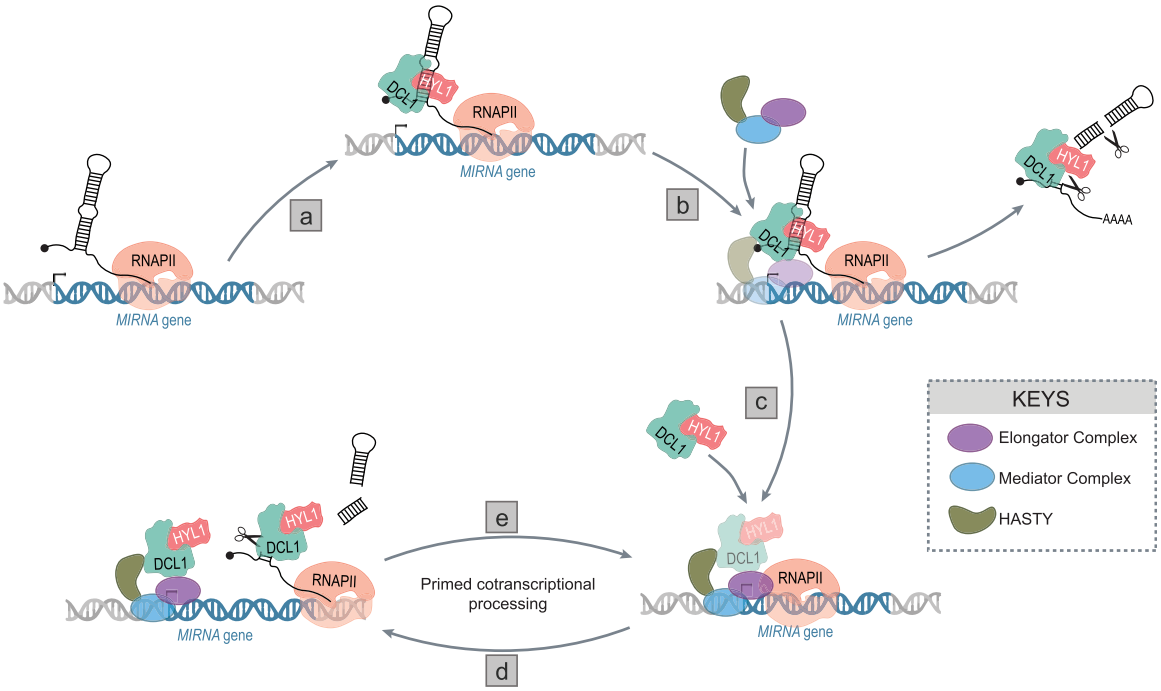


Figure 5 A "Processing Priming" hypothetical model for the recognition of *MIRNA* loci and initiation of cotranscriptional processing of pri-miRNAs. In a first round of Pol II-transcription, the dsRNA region of pri-miRNAs is recognized by the microprocessor (represented here by its core components DCL1 and HYL1) to initiate processing (a). DCL1 is then able to interact with the Elongation complex recruiting miRNA biogenesis-specific factors, such as HASTY, to the loci (b). The Elongator and Mediator complexes, along with HASTY and other accessory proteins, will then act as scaffolds promoting the recruitment of DCL1, and the microprocessor, to the primed *MIRNA* loci (c). This association enables a quick interaction of the microprocessor to nascent pri-miRNAs and their cotranscriptional processing (d) in a process that once primed is likely to perpetuate (e).

enable proper chromatin accessibility promoting the recruitment of the microprocessor to *MIRNA* loci allowing cotranscriptional processing.

Since these structures are relatively common in the genome of Arabidopsis (Xu et al., 2017, 2020) the presence of R-loops is unlikely the full explanation for the specificity toward the recognition of *MIRNA* loci. One possibility is that initial recognition of the dsRNA stem–loop region of a nascent pri-miRNA by the processing complex promotes the association of accessory factors, such as HST or SE, to the loci reinforcing the subsequent recruitment of the miRNA biogenesis machinery in a "primer" like mechanism (Figure 5). If such a scenario is accurate, then the processing complex may associate with any locus in the genome as long as a stem–loop-like structure is transcribed. Thus, it will be important to explore whether other proteins of the core miRNA biogenesis complex, namely DCL1 or HYL1, can bind other non-*MIRNA* loci in the genome.

Unlike other cotranscriptional RNA processing events, such as splicing or RNA editing, which can occur continuously as transcription progresses, the biogenesis of miRNA requires a substantial completion of transcription before it can happen. At least the entire stem–loop region of an miRNA precursor must be transcribed and correctly folded before DCL1 can recognize and slice it, giving the whole process a tight temporal window before transcription ends and the precursor is released. This is particularly relevant for plant miRNAs, as many precursors are rather long RNA molecules and contain complex secondary structures. The observation that changes in temperature and Pol II speed shift the frequency of cotranscriptional miRNA processing supports the idea of a tight temporal window (Gonzalo et al., 2022). Thus, it is likely that Pol II elongation speed, transcription termination, and cotranscriptional RNA folding play critical roles in modulating the temporal window required for cotranscriptional miRNA biogenesis.

The evidence suggests that coupling transcription and miRNA biogenesis can provide neo-functionalization to mature miRNAs. This turns the study of Pol II dynamics and transcription termination during miRNA biogenesis into essential topics. Surprisingly, it was recently shown that the assembly of the microprocessor at *MIRNA* loci assisted by PRE-MRNA PROCESSING PROTEIN 40 (PRP40), a U1 snRNP auxiliary protein, is not only important to trigger the cotranscriptional processing of miRNAs but also to promote Pol II-mediated transcription of miRNA genes (Stepien et al., 2022). This observation indicates that the crosstalk between the microprocessor and Pol II is bidirectional, opening the door to numerous regulatory scenarios.

Conclusions

RNA maturation is commonly represented as a stepwise process starting at transcription and mostly occurring in the nucleoplasm. This oversimplified view fails to consider the

complexity of the chromatin structure and its epigenetic modifications and ignores the crosstalk between the RNA polymerases and the complexes carrying out the different steps during RNA maturation. Within the concept of molecular crowding, caused by packing meters of linear DNA inside the nucleus, the coupling of RNA transcription and maturation presents an energetically and entropically favorable scenario. It is already an exciting notion for human cells, but plants such as *Paris japonica*, with a genome approximately 50 times larger than humans, take this concept to the extreme. In the future, our challenge resides in studying RNA maturation as a process associated with transcription while also considering that a nucleoplasm counterpart coexists. As a matter of fact, we should not only keep an eye on the interactions of each specific machineries with the RNA polymerases but also on their interactions with each other. Understanding whether cotranscriptional processing and modifications of specific RNAs may lead to differential functions compared to RNA maturated by canonical post-transcriptional mechanisms is at the frontier of our knowledge and a topic that we need to tackle in the near future. The current sequencing technology and the development of novel methodologies to sequence and characterize nascent RNAs have put us at the gate of exciting and unexpected discoveries.

Acknowledgments

We would like to thank ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Argentina) and CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) for funding to P.A.M. and E.P., and UNL (Universidad Nacional del Litoral) to P.A.M. Sebastian Marquardt (S.M.) acknowledges the funding from the Novo Nordisk Foundation NNF15OC0014202, NNF21OC0066776, and NNF19OC0057485 and Danmarks Frie Forskningsfond – Natur og Univers (0135-00094B). We would like to apologize to all those colleagues whose works could not be cited owing to space constraints.

Funding

This project received support from the European Research Council under the European Union's Horizon 2020 Research and Innovation Program StG2017-757411 to S.M.

Conflict of interest statement. None declared.

References

Adelman K, Marr MT, Werner J, Saunders A, Ni Z, Andrulis ED, Lis JT (2005) Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. Mol Cell **17**: 103–112

An D, Ma Q, Yan W, Zhou W, Liu G, Zhang P (2016) Divergent regulation of CBF regulon on cold tolerance and plant phenotype in cassava overexpressing Arabidopsis CBF3 gene. Front Plant Sci **7**: 1866

Arribas-Hernandez L, Rennie S, Koster T, Porcelli C, Lewinski M, Staiger D, Andersson R, Brodersen P (2021a) Principles of mRNA targeting via the Arabidopsis m(6)A-binding protein ECT2. eLife **10**: e72375

Arribas-Hernandez L, Rennie S, Schon M, Porcelli C, Enugutti B, Andersson R, Nodine MD, Brodersen P (2021b) The YTHDF proteins ECT2 and ECT3 bind largely overlapping target sets and influence target mRNA abundance, not alternative polyadenylation. eLife **10**: e72377

Andersen PR, Tirian L, Vunjak M, Brennecke J (2017) A heterochromatin-dependent transcription machinery drives piRNA expression. Nature **549**: 54–59

Anderson SJ, Kramer MC, Gosai SJ, Yu X, Vandivier LE, Nelson ADL, Anderson ZD, Beilstein MA, Fray RG, Lyons E, et al (2018) N(6)-methyladenosine inhibits local ribonucleolytic cleavage to stabilize mRNAs in Arabidopsis. Cell Rep **25**: 1146–1157 e1143

Andersson R, Sandelin A (2020) Determinants of enhancer and promoter activities of regulatory elements. Nat Rev Genet **21**: 71–87

Antosz W, Pfab A, Ehrnsberger HF, Holzinger P, Kollen K, Mortensen SA, Bruckmann A, Schubert T, Langst G, Griesenbeck J, et al. (2017) The composition of the Arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors. Plant Cell **29**: 854–870

Ariel F, Jegu T, Latrasse D, Romero-Barrios N, Christ A, Benhamed M, Crespi M (2014) Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. Mol Cell **55**: 383–396

Ariel FD, Manavella PA (2021) When junk DNA turns functional: transposon-derived noncoding RNAs in plants. J Exp Bot **72**: 4132–4143

Aslanzadeh V, Huang Y, Sanguinetti G, Beggs JD (2018) Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. Genome Res **28**: 203–213

Auboeuf D, Dowhan DH, Kang YK, Larkin K, Lee JW, Berget SM, O'Malley BW (2004) Differential recruitment of nuclear receptor coactivators may determine alternative RNA splice site choice in target genes. Proc Natl Acad Sci USA **101**: 2270–2274

Bedi K, Magnuson BR, Narayanan I, Paulsen M, Wilson TE, Ljungman M (2021) Cotranscriptional splicing efficiencies differ within genes and between cell types. RNA **27**: 829–840

Bentley DL (2014) Coupling mRNA processing with transcription in time and space. Nat Rev Genet **15**: 163–175

Berget SM, Moore C, Sharp PA (1977) Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci USA **74**: 3171–3175

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev **2**: 754–765

Bhat SS, Bielewicz D, Gulanicz T, Bodi Z, Yu X, Anderson SJ, Szewc L, Bajczyk M, Dolata J, Grzelak N, et al. (2020) mRNA adenosine methylase (MTA) deposits m(6)A on pri-miRNAs to modulate miRNA biogenesis in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **117**: 21785–21795

Bielewicz D, Kalak M, Kalyna M, Windels D, Barta A, Vazquez F, Szweykowska-Kulinska Z, Jarmolowski A (2013) Introns of plant pri-miRNAs enhance miRNA biogenesis. EMBO Rep **14**: 622–628

Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem **72**: 291–336

Boothby TC, Zipper RS, van der Weele CM, Wolniak SM (2013) Removal of retained introns regulates translation in the rapidly developing gametophyte of Marsilea vestita. Dev Cell **24**: 517–529

Boutz PL, Bhutkar A, Sharp PA (2015) Detained introns are a novel, widespread class of post-transcriptionally spliced introns. Genes Dev **29**: 63–80

Braberg H, Jin H, Moehle EA, Chan YA, Wang S, Shales M, Benschop JJ, Morris JH, Qiu C, Hu F, et al. (2013) From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. Cell **154**: 775–788

Bunch H, Zheng X, Burkholder A, Dillon ST, Motola S, Birrane G, Ebmeier CC, Levine S, Fargo D, Hu G, et al. (2014) TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release. Nat Struct Mol Biol **21**: 876–883

Buratowski S (2009) Progression through the RNA polymerase II CTD cycle. Mol Cell **36**: 541–546

Cambiagno DA, Giudicatti AJ, Arce AL, Gagliardi D, Li L, Yuan W, Lundberg DS, Weigel D, Manavella PA (2021) HASTY modulates miRNA biogenesis by linking pri-miRNA transcription and processing. Mol Plant **14**: 426–439

Cavalli G, Heard E (2019) Advances in epigenetics link genetics to the environment and disease. Nature **571**: 489–499

Chedin F (2016) Nascent connections: R-loops and chromatin patterning. Trends Genet **32**: 828–838

Chen T, Cui P, Xiong L (2015) The RNA-binding protein HOS5 and serine/arginine-rich proteins RS40 and RS41 participate in miRNA biogenesis in Arabidopsis. Nucleic Acids Res **43**: 8283–8298

Chow LT, Gelinas RE, Broker TR, Roberts RJ (1977) An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell **12**: 1–8

Corden JL (2013) RNA polymerase II C-terminal domain: tethering transcription to transcript and template. Chem Rev **113**: 8423–8455

Core L, Adelman K (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. Genes Dev **33**: 960–982

Cramer P (2019) Organization and regulation of gene transcription. Nature **573**: 45–54

Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA (2014) miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. Nature **508**: 411–415

Crevillen P, Sonmez C, Wu Z, Dean C (2013) A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J **32**: 140–148

Cyrek M, Fedak H, Ciesielski A, Guo Y, Sliwa A, Brzezniak L, Krzyczmonik K, Pietras Z, Kaczanowski S, Liu F, et al. (2016) Seed dormancy in Arabidopsis is controlled by alternative polyadenylation of DOG1. Plant Physiol **170**: 947–955

David R, Burgess A, Parker B, Li J, Pulsford K, Sibbritt T, Preiss T, Searle IR (2017) Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. Plant Cell **29**: 445–460

De Almeida C, Scheer H, Zuber H, Gagliardi D (2018) RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. Wiley Interdiscip Rev RNA **9**: e1440

de la Mata M, Kornblihtt AR (2006) RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat Struct Mol Biol **13**: 973–980

de la Mata M, Lafaille C, Kornblihtt AR (2010) First come, first served revisited: factors affecting the same alternative splicing event have different effects on the relative rates of intron removal. RNA **16**: 904–912

de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR (2003) A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell **12**: 525–532

de Lorenzo L, Sorenson R, Bailey-Serres J, Hunt AG (2017) Noncanonical alternative polyadenylation contributes to gene regulation in response to hypoxia. Plant Cell **29**: 1262–1277

Deutsch M, Long M (1999) Intron-exon structures of eukaryotic model organisms. Nucleic Acids Res **27**: 3219–3228

Dietrich MA, Prenger JP, Guilfoyle TJ (1990) Analysis of the genes encoding the largest subunit of RNA polymerase II in Arabidopsis and soybean. Plant Mol Biol **15**: 207–223

Dolata J, Guo Y, Kolowerzo A, Smolinski D, Brzyzek G, Jarmolowski A, Swiezewski S (2015) NTR1 is required for transcription elongation checkpoints at alternative exons in Arabidopsis. EMBO J **34**: 544–558

du Mee DJM, Ivanov M, Parker JP, Buratowski S, Marquardt S (2018) Efficient termination of nuclear lncRNA transcription promotes mitochondrial genome maintenance. eLife **7**: e31989

Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, et al. (2015) Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. EMBO J **34**: 581–592

Dvinge H (2018) Regulation of alternative mRNA splicing: old players and new perspectives. FEBS Lett **592**: 2987–3006

Fang X, Cui Y, Li Y, Qi Y (2015) Transcription and processing of primary microRNAs are coupled by Elongator complex in Arabidopsis. Nat Plants **1**: 15075

Fong N, Kim H, Zhou Y, Ji X, Qiu J, Saldi T, Diener K, Jones K, Fu XD, Bentley DL (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. Genes Dev **28**: 2663–2676

Francette AM, Tripplehorn SA, Arndt KM (2021) The paf1 complex: a keystone of nuclear regulation operating at the interface of transcription and chromatin. J Mol Biol **433**: 166979

Gagliardi D, Manavella PA (2020) Short-range regulatory chromatin loops in plants. New Phytol **228**: 466–471

Gagliardi D, Cambiagno DA, Arce AL, Tomassi AH, Giacomelli JI, Ariel FD, Manavella PA (2019) Dynamic regulation of chromatin topology and transcription by inverted repeat-derived small RNAs in sunflower. Proc Natl Acad Sci USA **116**: 17578–17583

Gilbert W (1978) Why genes in pieces? Nature **271**: 501

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. Plant Physiol **124**: 1854–1865

Godoy Herz MA, Kubaczka MG, Brzyzek G, Servi L, Krzyszton M, Simpson C, Brown J, Swiezewski S, Petrillo E, Kornblihtt AR (2019) Light regulates plant alternative splicing through the control of transcriptional elongation. Mol Cell **73**: 1066–1074 e1063

Gonzalo L, Tossolini I, Gulanicz T, Cambiagno DA, Kasprowicz-Maluski A, Smolinski DJ, Mammarella MF, Ariel FD, Marquardt S, Szweykowska-Kulinska Z, et al. (2022) R-loops at microRNA encoding loci promote cotranscriptional processing of pri-miRNAs in plants. Nat Plants **8**: 402–418

Gowthaman U, Garcia-Pichardo D, Jin Y, Schwarz I, Marquardt S (2020) DNA processing in the context of noncoding transcription. Trends Biochem Sci **45**: 1009–1021

Gowthaman U, Ivanov M, Schwarz I, Patel HP, Muller NA, Garcia-Pichardo D, Lenstra TL, Marquardt S (2021) The Hda1 histone deacetylase limits divergent noncoding transcription and restricts transcription initiation frequency. EMBO J **40**: e108903

Guo L, Cao X, Liu Y, Li J, Li Y, Li D, Zhang K, Gao C, Dong A, Liu X (2018) A chromatin loop represses WUSCHEL expression in Arabidopsis. Plant J **94**: 1083–1097

Guo YE, Manteiga JC, Henninger JE, Sabari BR, Dall'Agnese A, Hannett NM, Spille JH, Afeyan LK, Zamudio AV, Shrinivas K, et al. (2019) Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature **572**: 543–548

Haag JR, Pikaard CS (2011) Multisubunit RNA polymerases IV and V: purveyors of noncoding RNA for plant gene silencing. Nat Rev Mol Cell Biol **12**: 483–492

Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, Pikaard CS (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol Cell **48**: 811–818

Haberle V, Stark A (2018) Eukaryotic core promoters and the functional basis of transcription initiation. Nat Rev Mol Cell Biol **19**: 621–637

Hawkins JD (1988) A survey on intron and exon lengths. Nucleic Acids Res **16**: 9893–9908

Hazelbaker DZ, Marquardt S, Wlotzka W, Buratowski S (2013) Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. Mol Cell **49**: 55–66

Henriques R, Mas P (2013) Chromatin remodeling and alternative splicing: pre- and post-transcriptional regulation of the Arabidopsis circadian clock. Semin Cell Dev Biol **24**: 399–406

Herzel L, Ottoz DSM, Alpert T, Neugebauer KM (2017) Splicing and transcription touch base: cotranscriptional spliceosome assembly and function. Nat Rev Mol Cell Biol **18**: 637–650

Hornyik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. Dev Cell **18**: 203–213

Hu Q, Kim EJ, Feng J, Grant GR, Heller EA (2017) Histone post-translational modifications predict specific alternative exon subtypes in mammalian brain. PLoS Comput Biol **13**: e1005602

Hung YH, Slotkin RK (2021) The initiation of RNA interference (RNAi) in plants. Curr Opin Plant Biol **61**: 102014

Isbel L, Grand RS, Schubeler D (2022) Generating specificity in genome regulation through transcription factor sensitivity to chromatin. Nat Rev Genet https://doi.org/10.1038/s41576-022-00512-6

Ivanov M, Sandelin A, Marquardt S (2021) Publisher correction to: trancriptomeReconstructoR: datadriven annotation of complex transcriptomes. BMC Bioinformatics **22**: 370

Jabre I, Reddy ASN, Kalyna M, Chaudhary S, Khokhar W, Byrne LJ, Wilson CM, Syed NH (2019) Does cotranscriptional regulation of alternative splicing mediate plant stress responses? Nucleic Acids Res **47**: 2716–2726

Jia T, Zhang B, You C, Zhang Y, Zeng L, Li S, Johnson KCM, Yu B, Li X, Chen X (2017) The Arabidopsis MOS4-associated complex promotes microRNA biogenesis and precursor messenger RNA splicing. Plant Cell **29**: 2626–2643

Jimeno-Gonzalez S, Payan-Bravo L, Munoz-Cabello AM, Guijo M, Gutierrez G, Prado F, Reyes JC (2015) Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing. Proc Natl Acad Sci USA **112**: 14840–14845

Kamieniarz-Gdula K, Proudfoot NJ (2019) Transcriptional control by premature termination: a forgotten mechanism. Trends Genet **35**: 553–564

Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, et al. (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell **123**: 593–605

Khodor YL, Rodriguez J, Abruzzi KC, Tang CH, Marr MT 2nd, Rosbash M (2011) Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes Dev **25**: 2502–2512

Kim A, Wang GG (2021) R-loop and its functions at the regulatory interfaces between transcription and (epi)genome. Biochim Biophys Acta Gene Regul Mech **1864**: 194750

Kim DH, Sung S (2017) Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. Dev Cell **40**: 302–312 e304

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

Kim T, Xu Z, Clauder-Munster S, Steinmetz LM, Buratowski S (2012) Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. Cell **150**: 1158–1169

Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X (2011) The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. EMBO J **30**: 814–822

Kindgren P, Ivanov M, Marquardt S (2020) Native elongation transcript sequencing reveals temperature dependent dynamics of nascent Pol II transcription in Arabidopsis. Nucleic Acids Res **48**: 2332–2347

Kindgren P, Ard R, Ivanov M, Marquardt S (2018) Transcriptional read-through of the long noncoding RNA SVALKA governs plant cold acclimation. Nat Commun **9**: 4561

Kitsios G, Alexiou KG, Bush M, Shaw P, Doonan JH (2008) A cyclin-dependent protein kinase, CDKC2, colocalizes with and modulates the distribution of spliceosomal components in Arabidopsis. Plant J **54**: 220–235

Knop K, Stepien A, Barciszewska-Pacak M, Taube M, Bielewicz D, Michalak M, Borst JW, Jarmolowski A, Szweykowska-Kulinska Z (2017) Active 5′ splice sites regulate the biogenesis efficiency of Arabidopsis microRNAs derived from intron-containing genes. Nucleic Acids Res **45**: 2757–2775

Komarnitsky P, Cho EJ, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev **14**: 2452–2460

Kornblihtt AR, Schor IE, Allo M, Dujardin G, Petrillo E, Munoz MJ (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol **14**: 153–165

Koster T, Meyer K, Weinholdt C, Smith LM, Lummer M, Speth C, Grosse I, Weigel D, Staiger D (2014) Regulation of pri-miRNA processing by the hnRNP-like protein AtGRP7 in Arabidopsis. Nucleic Acids Res **42**: 9925–9936

Kowalik KM, Shimada Y, Flury V, Stadler MB, Batki J, Buhler M (2015) The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. Nature **520**: 248–252

Kumar S, Mohapatra T (2021) Deciphering epitranscriptome: modification of mRNA bases provides a new perspective for post-transcriptional regulation of gene expression. Front Cell Dev Biol **9**: 628415

Lahmy S, Pontier D, Bies-Etheve N, Laudie M, Feng S, Jobet E, Hale CJ, Cooke R, Hakimi MA, Angelov D, et al. (2016) Evidence for ARGONAUTE4-DNA interactions in RNA-directed DNA methylation in plants. Genes Dev **30**: 2565–2570

Lai X, Blanc-Mathieu R, GrandVuillemin L, Huang Y, Stigliani A, Lucas J, Thevenon E, Loue-Manifel J, Turchi L, Daher H, et al. (2021) The LEAFY floral regulator displays pioneer transcription factor properties. Mol Plant **14**: 829–837

Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann JU, Ratsch G, Weigel D (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **105**: 8795–8800

Leng X, Thomas Q, Rasmussen SH, Marquardt S (2020a) A genomic positioning system for plant pol II transcription. Trends Plant Sci **25**: 744–764

Leng X, Ivanov M, Kindgren P, Malik I, Thieffry A, Brodersen P, Sandelin A, Kaplan CD, Marquardt S (2020b) Organismal benefits of transcription speed control at gene boundaries. EMBO Rep **21**: e49315

Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. Cell **128**: 707–719

Li P, Tao Z, Dean C (2015) Phenotypic evolution through variation in splicing of the noncoding RNA COOLAIR. Genes Dev **29**: 696–701

Li S, Wang Y, Zhao Y, Zhao X, Chen X, Gong Z (2020) Global cotranscriptional splicing in Arabidopsis and the correlation with splicing regulation in mature RNAs. Mol Plant **13**: 266–277

Liu C, Teo ZW, Bi Y, Song S, Xi W, Yang X, Yin Z, Yu H (2013) A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice. Dev Cell **24**: 612–622

Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3′ processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science **327**: 94–97

Liu W, Duttke SH, Hetzel J, Groth M, Feng S, Gallego-Bartolome J, Zhong Z, Kuo HY, Wang Z, Zhai J, et al. (2018) RNA-directed DNA methylation involves cotranscriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. Nat Plants **4**: 181–188

Liu Y, Geyer R, van Zanten M, Carles A, Li Y, Horold A, van Nocker S, Soppe WJ (2011) Identification of the Arabidopsis REDUCED DORMANCY 2 gene uncovers a role for the polymerase associated factor 1 complex in seed dormancy. PLoS One **6**: e22241

Loffer A, Singh J, Fukudome A, Mishra V, Wang F, Pikaard CS (2022) A DCL3 dicing code within Pol IV-RDR2 transcripts diversifies the siRNA pool guiding RNA-directed DNA methylation. eLife **11**: e73260

Louloupi A, Ntini E, Conrad T, Orom UAV (2018) Transient N-6-methyladenosine transcriptome sequencing reveals a regulatory role of m6A in splicing efficiency. Cell Rep **23**: 3429–3437

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature **389**: 251–260

Luo GZ, MacQueen A, Zheng G, Duan H, Dore LC, Lu Z, Liu J, Chen K, Jia G, Bergelson J, et al. (2014) Unique features of the m6A methylome in *Arabidopsis thaliana*. Nat Commun **5**: 5630

Manavella PA, Hagmann J, Ott F, Laubinger S, Franz M, Macek B, Weigel D (2012) Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell **151**: 859–870

Marquardt S, Raitskin O, Wu Z, Liu F, Sun Q, Dean C (2014a) Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. Mol Cell **54**: 156–165

Marquardt S, Escalante-Chong R, Pho N, Wang J, Churchman LS, Springer M, Buratowski S (2014b) A chromatin-based mechanism for limiting divergent noncoding transcription. Cell **157**: 1712–1723

Marquez Y, Brown JW, Simpson C, Barta A, Kalyna M (2012) Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome Res **22**: 1184–1195

Marquez Y, Hopfler M, Ayatollahi Z, Barta A, Kalyna M (2015) Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. Genome Res **25**: 995–1007

Martin G, Marquez Y, Mantica F, Duque P, Irimia M (2021) Alternative splicing landscapes in *Arabidopsis thaliana* across tissues and stress conditions highlight major functional differences with animals. Genome Biol **22**: 35

Martienssen R, Moazed D (2015) RNAi and heterochromatin assembly. Cold Spring Harb Perspect Biol **7**: a019323

Maslon MM, Braunschweig U, Aitken S, Mann AR, Kilanowski F, Hunter CJ, Blencowe BJ, Kornblihtt AR, Adams IR, Caceres JF (2019) A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. EMBO J **38**: e101244

Mencia R, Gonzalo L, Tossolini I, Manavella PA (2022) Keeping up with the miRNAs: current paradigms of the biogenesis pathway. J Exp Bot https://doi.org/10.1093/jxb/erac322

Mendel M, Delaney K, Pandey RR, Chen KM, Wenda JM, Vagbo CB, Steiner FA, Homolka D, Pillai RS (2021) Splice site m(6)A methylation prevents binding of U2AF35 to inhibit RNA splicing. Cell **184**: 3125–3142 e3125

Mishra V, Singh J, Wang F, Zhang Y, Fukudome A, Trinidad JC, Takagi Y, Pikaard CS (2021) Assembly of a dsRNA synthesizing complex: RNA-DEPENDENT RNA POLYMERASE 2 contacts the largest subunit of NUCLEAR RNA POLYMERASE IV. Proc Natl Acad Sci U S A **118**: e2019276118

Moretto F, Wood NE, Kelly G, Doncic A, van Werven FJ (2018) A regulatory circuit of two lncRNAs and a master regulator directs cell fate in yeast. Nat Commun **9**: 780

Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ (2008) Primary microRNA transcripts are processed cotranscriptionally. Nat Struct Mol Biol **15**: 902–909

Ner-Gaon H, Halachmi R, Savaldi-Goldstein S, Rubin E, Ophir R, Fluhr R (2004) Intron retention is a major phenomenon in alternative splicing in Arabidopsis. Plant J **39**: 877–885

Nielsen M, Ard R, Leng X, Ivanov M, Kindgren P, Pelechano V, Marquardt S (2019) Transcription-driven chromatin repression of Intragenic transcription start sites. PLoS Genet **15**: e1007969

Nojima T, Proudfoot NJ (2022) Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. Nat Rev Mol Cell Biol **23**: 389–406

Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, Carmo-Fonseca M, Proudfoot NJ (2015) Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. Cell **161**: 526–540

Oh S, Zhang H, Ludwig P, van Nocker S (2004) A mechanism related to the yeast transcriptional regulator Paf1c is required for expression of the Arabidopsis FLC/MAF MADS box gene family. Plant Cell **16**: 2940–2953

Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK (2016) Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. Genome Biol **17**: 170

Parker MT, Knop K, Zacharaki V, Sherwood AV, Tome D, Yu X, Martin PG, Beynon J, Michaels SD, Barton GJ, et al. (2021) Widespread premature transcription termination of *Arabidopsis thaliana* NLR genes by the spen protein FPA. eLife **10**: e65537

Parker MT, Knop K, Sherwood AV, Schurch NJ, Mackinnon K, Gould PD, Hall AJ, Barton GJ, Simpson GG (2020) Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. eLife **9**: e49658

Pelechano V, Wei W, Steinmetz LM (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature **497**: 127–131

Proudfoot NJ (2016) Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. Science **352**: aad9926

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. Cell **108**: 501–512

Quesada V, Macknight R, Dean C, Simpson GG (2003) Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. EMBO J **22**: 3142–3152

Raczynska KD, Stepien A, Kierzkowski D, Kalak M, Bajczyk M, McNicol J, Simpson CG, Szweykowska-Kulinska Z, Brown JW, Jarmolowski A (2014) The SERRATE protein is involved in alternative splicing in *Arabidopsis thaliana*. Nucleic Acids Res **42**: 1224–1244

Rodriguez CR, Cho EJ, Keogh MC, Moore CL, Greenleaf AL, Buratowski S (2000) Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. Mol Cell Biol **20**: 104–112

Rougvie AE, Lis JT (1988) The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell **54**: 795–804

Sasaki T, Lee TF, Liao WW, Naumann U, Liao JL, Eun C, Huang YY, Fu JL, Chen PY, Meyers BC, et al. (2014) Distinct and concurrent pathways of Pol II- and Pol IV-dependent siRNA biogenesis at a repetitive trans-silencer locus in *Arabidopsis thaliana*. Plant J **79**: 127–138

Schor IE, Gomez Acuna LI, Kornblihtt AR (2013) Coupling between transcription and alternative splicing. Cancer Treat Res **158**: 1–24

Sharp PA, Chakraborty AK, Henninger JE, Young RA (2022) RNA in formation and regulation of transcriptional condensates. RNA **28**: 52–57

Sigman MJ, Panda K, Kirchner R, McLain LL, Payne H, Peasari JR, Husbands AY, Slotkin RK, McCue AD (2021) An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. Nat Plants **7**: 1461–1474

Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ (2010) Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. Mol Cell **38**: 202–210

Singh J, Mishra V, Wang F, Huang HY, Pikaard CS (2019) Reaction mechanisms of Pol IV, RDR2, and DCL3 Drive RNA channeling in the siRNA-directed DNA methylation pathway. Mol Cell **75**: 576–589 e575

Skalska L, Beltran-Nebot M, Ule J, Jenner RG (2017) Regulatory feedback from nascent RNA to chromatin and transcription. Nat Rev Mol Cell Biol **18**: 331–337

Stepien A, Dolata J, Gulanicz T, Bielewicz D, Bajczyk M, Smolinski DJ, Szweykowska-Kulinska Z, Jarmolowski A (2022)

Chromatin-associated microprocessor assembly is regulated by the U1 snRNP auxiliary protein PRP40. Plant Cell **34**: 4920–4935

Szweykowska-Kulinska Z, Jarmolowski A, Vazquez F (2013) The crosstalk between plant microRNA biogenesis factors and the spliceosome. Plant Signal Behav **8**: e26955

Tan-Wong SM, Zaugg JB, Camblong J, Xu Z, Zhang DW, Mischo HE, Ansari AZ, Luscombe NM, Steinmetz LM, Proudfoot NJ (2012) Gene loops enhance transcriptional directionality. Science **338**: 671–675

Tellier M, Maudlin I, Murphy S (2020) Transcription and splicing: a two-way street. Wiley Interdiscip Rev RNA **11**: e1593

Thomas QA, Ard R, Liu J, Li B, Wang J, Pelechano V, Marquardt S (2020) Transcript isoform sequencing reveals widespread promoter-proximal transcriptional termination in Arabidopsis. Nat Commun **11**: 2589

Tomassi AH, Re DA, Romani F, Cambiagno DA, Gonzalo L, Moreno JE, Arce AL, Manavella PA (2020) The intrinsically disordered protein CARP9 bridges HYL1 to AGO1 in the nucleus to promote microRNA activity. Plant Physiol **184**: 316–329

Tsuzuki M, Sethuraman S, Coke AN, Rothi MH, Boyle AP, Wierzbicki AT (2020) Broad noncoding transcription suggests genome surveillance by RNA polymerase V. Proc Natl Acad Sci USA **117**: 30799–30804

Vandivier LE, Campos R, Kuksa PP, Silverman IM, Wang LS, Gregory BD (2015) Chemical modifications mark alternatively spliced and uncapped messenger RNAs in Arabidopsis. Plant Cell **27**: 3024–3037

Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL (2012) Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature **489**: 452–455

Wang F, Axtell MJ (2017) AGO4 is specifically required for heterochromatic siRNA accumulation at Pol V-dependent loci in *Arabidopsis thaliana*. Plant J **90**: 37–47

Wang L, Song X, Gu L, Li X, Cao S, Chu C, Cui X, Chen X, Cao X (2013) NOT2 proteins promote polymerase II-dependent transcription and interact with multiple MicroRNA biogenesis factors in Arabidopsis. Plant Cell **25**: 715–727

Wang S, Quan L, Li S, You C, Zhang Y, Gao L, Zeng L, Liu L, Qi Y, Mo B, et al. (2019) The PROTEIN PHOSPHATASE4 complex promotes transcription and processing of primary microRNAs in Arabidopsis. Plant Cell **31**: 486–501

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, et al. (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature **505**: 117–120

Wei W, Hennig BP, Wang J, Zhang Y, Piazza I, Pareja Sanchez Y, Chabbert CD, Adjalley SH, Steinmetz LM, Pelechano V (2019) Chromatin-sensitive cryptic promoters putatively drive expression of alternative protein isoforms in yeast. Genome Res **29**: 1974–1984

Wendte JM, Haag JR, Pontes OM, Singh J, Metcalf S, Pikaard CS (2019) The Pol IV largest subunit CTD quantitatively affects siRNA levels guiding RNA-directed DNA methylation. Nucleic Acids Res **47**: 9024–9036

West S, Gromak N, Proudfoot NJ (2004) Human 5′ –> 3′ exonuclease Xrn2 promotes transcription termination at cotranscriptional cleavage sites. Nature **432**: 522–525

Williamson L, Saponaro M, Boeing S, East P, Mitter R, Kantidakis T, Kelly GP, Lobley A, Walker J, Spencer-Dene B, et al. (2017) UV irradiation induces a noncoding RNA that functionally opposes the protein encoded by the same gene. Cell **168**: 843–855 e813

Workman JL, Kingston RE (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annu Rev Biochem **67**: 545–579

Wu Z, Fang X, Zhu D, Dean C (2020) Autonomous pathway: FLOWERING LOCUS C repression through an antisense-mediated chromatin-silencing mechanism. Plant Physiol **182**: 27–37

Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, Sun HY, Li A, Ping XL, Lai WY, et al. (2016) Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell **61**: 507–519

Xu C, Wu Z, Duan HC, Fang X, Jia G, Dean C (2021) R-loop resolution promotes cotranscriptional chromatin silencing. Nat Commun **12**: 1790

Xu W, Xu H, Li K, Fan Y, Liu Y, Yang X, Sun Q (2017) The R-loop is a common chromatin feature of the Arabidopsis genome. Nat Plants **3**: 704–714

Xu W, Li K, Li S, Hou Q, Zhang Y, Liu K, Sun Q (2020) The R-loop atlas of Arabidopsis development and responses to environmental stimuli. Plant Cell **32**: 888–903

Yang J, Cao Y, Ma L (2021) Cotranscriptional RNA processing in plants: exploring from the perspective of polyadenylation. Int J Mol Sci **22**: 3300

Yang L, Perrera V, Saplaoura E, Apelt F, Bahin M, Kramdi A, Olas J, Mueller-Roeber B, Sokolowska E, Zhang W, et al. (2019) m(5)C Methylation guides systemic transport of messenger RNA over graft junctions in plants. Curr Biol **29**: 2465–2476 e2465

Ye R, Chen Z, Lian B, Rowley MJ, Xia N, Chai J, Li Y, He XJ, Wierzbicki AT, Qi Y (2016) A dicer-independent route for biogenesis of siRNAs that direct DNA methylation in Arabidopsis. Mol Cell **61**: 222–235

Yu H, Tian C, Yu Y, Jiao Y (2016) Transcriptome survey of the contribution of alternative splicing to proteome diversity in *Arabidopsis thaliana*. Mol Plant **9**: 749–752

Yu Z, Hong L, Li QQ (2022) Signatures of mRNA alternative polyadenylation in Arabidopsis leaf development. Front Genet **13**: 863253

Zaret KS (2020) Pioneer transcription factors initiating gene network changes. Annu Rev Genet **54**: 367–385

Zhang R, Kuo R, Coulter M, Calixto CPG, Entizne JC, Guo W, Marquez Y, Milne L, Riegler S, Matsui A, et al. (2022) A high-resolution single-molecule sequencing-based Arabidopsis transcriptome using novel methods of Iso-seq analysis. Genome Biol **23**: 149

Zhang S, Li R, Zhang L, Chen S, Xie M, Yang L, Xia Y, Foyer CH, Zhao Z, Lam HM (2020) New insights into Arabidopsis transcriptome complexity revealed by direct sequencing of native RNAs. Nucleic Acids Res **48**: 7700–7711

Zhang S, Liu Y, Yu B (2014) PRL1, an RNA-binding protein, positively regulates the accumulation of miRNAs and siRNAs in Arabidopsis. PLoS Genet **10**: e1004841

Zhang S, Dou Y, Li S, Ren G, Chevalier D, Zhang C, Yu B (2018) DAWDLE interacts with DICER-LIKE proteins to mediate small RNA biogenesis. Plant Physiol **177**: 1142–1151

Zhang T, Marand AP, Jiang J (2016) PlantDHS: a database for DNase I hypersensitive sites in plants. Nucleic Acids Res **44**: D1148–D1153

Zhao Y, Zhu P, Hepworth J, Bloomer R, Antoniou-Kourounioti RL, Doughty J, Heckmann A, Xu C, Yang H, Dean C (2021) Natural temperature fluctuations promote COOLAIR regulation of FLC. Genes Dev **35**: 888–898

Zhu D, Mao F, Tian Y, Lin X, Gu L, Gu H, Qu LJ, Wu Y, Wu Z (2020) The features and regulation of cotranscriptional splicing in Arabidopsis. Mol Plant **13**: 278–294