Classification of Transcription Boundary-Associated RNAs (TBARs) in Animals and Plants

Dongliang Yu†, Xiaoxia Ma†, Ziwei Zuo, Huizhong Wang and Yijun Meng*

College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China

There is increasing evidence suggesting the contribution of non-coding RNAs (ncRNAs) to the phenotypic and physiological complexity of organisms. A novel ncRNA species has been identified near the transcription boundaries of protein-coding genes in eukaryotes, bacteria, and archaea. This review provides a detailed description of these transcription boundary-associated RNAs (TBARs), including their classification. Based on their genomic distribution, TBARs are divided into two major groups: promoter-associated RNAs (PARs) and terminus-associated RNAs (TARs). Depending on the sequence length, each group is further classified into long RNA species (>200 nt) and small RNA species (<200 nt). According to these rules of TBAR classification, divergent ncRNAs with confusing nomenclatures, such as promoter upstream transcripts (PROMPTs), upstream antisense RNAs (uaRNAs), stable unannotated transcripts (SUTs), cryptic unstable transcripts (CUTs), transcription start site-associated RNAs (TSSaRNAs), transcription initiation RNAs (tIRNAs), and transcription termination site-associated RNAs (TTSaRNAs), were assigned to specific classes. Although the biogenesis pathways of PARs and TARs have not yet been clearly elucidated, previous studies indicate that some of the PARs have originated either through divergent transcription or via RNA polymerase pausing. Intriguing findings regarding the functional implications of the TBARs such as the long-range “gene looping” model, which explains their role in the transcriptional regulation of protein-coding genes, are also discussed. Altogether, this review provides a comprehensive overview of the current research status of TBARs, which will promote further investigations in this research area.

Keywords: transcription boundary-associated RNAs (TBARs), promoter-associated RNAs (PARs), terminus-associated RNAs (TARs), divergent transcription, transcription start site (TSS), transcription termination site (TTS), polymerase (Pol) pausing, gene looping

INTRODUCTION

In the initial phase of genome sequencing, annotated protein-coding genes were thought to be a major contributor to the development of phenotypic and physiological complexity of organisms. However, this viewpoint was challenged soon afterwards. The application of high-throughput sequencing (HTS) in transcriptomics studies uncovered the intriguing phenomenon in which transcription signals were detected across the entire genomes (Encode Project Consortium et al., 2007). Notably, most of these signals originated from non-coding loci but not from...
protein-coding genes. Subsequent studies have revealed a positive correlation between the diversity of non-coding RNAs (ncRNAs) and the complexity of organisms (Amaral and Mattick, 2008; Amaral et al., 2008). Considering the emerging biological importance of ncRNAs, continuous research efforts have led to the identification and characterization of ncRNAs in multiple organisms. Although a large proportion of ncRNAs resides within intergenic regions, some reside within protein-coding genes, such as antisense ncRNAs (He et al., 2008) and intronic ncRNAs (Brown et al., 2008; Meng et al., 2011; Meng and Shao, 2012). Additionally, some of the ncRNAs have been identified near the promoter regions or the transcription termini of protein-coding genes (Gingeras, 2007; Kapranov et al., 2007).

In this review, we provide a brief summary of the discovery of promoter-associated RNAs (PARs) in budding yeast (Saccharomyces cerevisiae) (Davis and Ares, 2006; Neil et al., 2009; Xu et al., 2009), fruit fly (Drosophila melanogaster), chicken (Gallus gallus) (Taft et al., 2009a), mouse (Mus musculus) (Kapranov et al., 2007; Seila et al., 2008; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009; Flynn et al., 2011), human (Homo sapiens) (Han et al., 2007; Kapranov et al., 2007; Preker et al., 2008; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009; Pickrell et al., 2010), Arabidopsis thaliana (Chekanova et al., 2007; Wang et al., 2011; Ma et al., 2017), bacteria (Yus et al., 2012), and archaea (Zaramela et al., 2014). Additionally, the discovery of terminus-associated RNAs (TARs) in human (Kapranov et al., 2007, 2010; Taft et al., 2009a; Yue et al., 2010; Valen et al., 2011; Schwalb et al., 2016), mouse (Kapranov et al., 2007), fission yeast (Schizosaccharomyces pombe), and Arabidopsis (Ma et al., 2017) has also been introduced. Since promoters and terminators define the transcription boundaries of protein-coding genes, PARs and TARs are referred to as transcription boundary-associated RNAs (TBARs). Additionally, we summarize the biogenesis pathways of some of the TBARs and a few mechanistic models of TBAR biogenesis including divergent transcription and RNA polymerase (Pol) pausing. We also summarize recent progress in the functional studies conducted on TBARs, followed by in-depth discussions. Interestingly, PARs have been functionally implicated in transcription termination, whereas TARs are potentially involved in transcription initiation. The interactive relationship between transcription initiation and termination has been depicted using long-range regulatory models, such as “gene looping” (Lykke-Andersen et al., 2011; Andersen et al., 2013). Finally, we propose a preliminary nomenclature system for TBARs according to specific classification criteria and emphasize the importance of a uniform annotation system for TBAR research.

**PARs IN ANIMALS: DISCOVERY, BIOGENESIS, AND FUNCTIONS**

In eukaryotes, the transcription of protein-coding genes includes two primary steps: initiation and elongation. Each step is under strict surveillance. However, each time transcription is initiated from the transcription start site (TSS) of a protein-coding gene, the final output is not always a full-length coding sequence. This is because various kinds of genetic and epigenetic factors, such as binding sites of RNA Pol II or transcription factors (TFs) and chromatin modifications, interrupt transcription initiation and/or elongation (Li et al., 2007; Razin et al., 2011). Premature termination of transcription is a widespread phenomenon that results in aberrant transcripts. As a result of frequent transcription pausing, diverse ncRNAs have been discovered near the promoter regions. Although such ncRNAs have been referred to by different names, here they are referred to as PARs.

In a pioneering study, Core et al. (2008) employed global run-on sequencing (GRO-Seq) to investigate the density of transcriptionally engaged RNA polymerase enzymes on a genome-wide scale. This study identified a novel species of PARs in human lung fibroblasts. However, because this study used an older generation sequencing platform with short read-length (maximum 33 nt), data on the length of PARs was not provided. These PARs are enriched within ~250 bp to +50 bp of the TSSs of certain protein-coding genes. Because GRO-Seq is capable of evaluating promoter-proximal pausing on all genes, it has been suggested that these PARs are derived from RNA Pol II-dependent divergent transcription and promoter-proximal pausing (Core et al., 2008). Subsequent studies have uncovered PARs in different species. Because of recent advances in sequencing platforms, more information is available on the sequence features of PARs, especially their length distribution. In this review, we propose that, according to their sequence length, PARs should to be divided into two classes: promoter-associated long RNAs (PALRs; >200 nt) and promoter-associated small RNAs (PASRs; <200 nt).

Promoter-associated small RNAs have been reported in human and mouse (Kapranov et al., 2007; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009). TSS-associated RNAs (TSSaRNAs) reported in mouse (Seila et al., 2008) and transcription initiation RNAs (tiRNAs) in human, mouse, chicken and Drosophila (Taft et al., 2009a) have also been categorized as PASRs. Based on these studies, it can be concluded that weak expression is a common feature of PASRs. A number of PARs are exosome-specific substrates with short half-lives (Wyers et al., 2005; Davis and Ares, 2006; Chekanova et al., 2007; Preker et al., 2008). The divergent distribution of many PARs near the TSSs is correlated with bidirectional transcription activity of RNA Pol II. Functional studies have connected some of the PARs with transcriptional activation or repression of their target genes (Li et al., 2006; Janowski et al., 2007; Morris et al., 2008; Turunen et al., 2009; Huang et al., 2010; Kassab et al., 2015; Li et al., 2016; Uesaka et al., 2017), indicating that transcription can be modulated not only during elongation but also at initiation (Saunders et al., 2006; Muse et al., 2007; Zeitlinger et al., 2007; Core et al., 2008).

In the following sections, we describe the previously reported PARs in animals and their biogenesis pathways and biological functions.
Results of tiling microarray analysis revealed 5 the identification of short-lived PARs using oligodT primers. (Preker et al., 2008, 2011) used human HeLa cells without exosome activity for detection. To overcome this limitation, Preker et al. (2008, 2011), upstream antisense RNAs (uaRNAs) in mouse (Flynn et al., 2011), stable unannotated transcripts (SUTs) and cryptic unstable transcripts (CUTs) in yeast (Wyers et al., 2005; Neil et al., 2009; Xu et al., 2009), and other unnamed PALRs in human and mouse (Kapranov et al., 2007; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009). Till date, only a few studies have been conducted on the biogenesis pathways and modes of action of PALRs. Thus, it is unclear whether these PALRs, with distinct nomenclatures, belong to the same or different ncRNA species. It is well known that divergent transcription of active promoters is the major mechanism involved in the biogenesis of PALRs (Sigova et al., 2013). Among the above-mentioned PALRs, PROMPTs are relatively well studied, especially for their sequence characteristics, biogenesis and biological roles. Therefore, we describe the PROMPTs in detail in the sections below.

Discovery of PROMPTs
Promoter-associated RNAs have short half-lives, as they undergo rapid exosome-mediated degradation, which interferes with their detection. To overcome this limitation, Preker et al. (2008, 2011) used human HeLa cells without exosome activity for the identification of short-lived PARs using oligodT primers. Results of tiling microarray analysis revealed 5′ capped and 3′ adenylated transcripts upstream of the TSSs of transcriptionally active genes (Preker et al., 2008, 2011). These transcripts were several hundred nucleotides in length and were named as PROMPTs.

Unlike PARs discovered by Core et al. (2008), PROMPTs originate from the region within −2,500 bp to −500 bp upstream of the TSS. Moreover, the transcription of PROMPTs can proceed in both directions (Preker et al., 2008). It is widely accepted that divergent transcription is an inherent feature of most of the active promoters (Beck and Warren, 1988; Neil et al., 2009; Xu et al., 2009). Moreover, the production of PROMPTs is highly dependent on activity of promoters of neighboring genes (Preker et al., 2008). Thus, the bidirectional transcription of PROMPTs reflects the universality of divergent transcription of highly active genes, which links bidirectional transcription to gene activity (Core et al., 2008; Nechaev and Adelman, 2011).

Functions of PROMPTs
Functional studies suggest that bidirectional transcription plays an important role in improving the accessibility of chromatin regions for the binding of TFs (Gilchrist et al., 2008; Seila et al., 2008). Notably, some of the divergently transcribed PARs are involved in Argonaute (AGO)-dependent (Han et al., 2007) or -independent (Wang et al., 2008) gene silencing pathways, indicating a potential regulatory role of PROMPTs in gene transcription. The SR proteins facilitate the function of promoter-proximal nascent RNA in transcription pause release (Ji et al., 2013). Additionally, some of the PROMPTs are enriched within the promoter regions with a high CpG content, and these PROMPTs affect the DNA methylation density of the promoters (Preker et al., 2008).

TABLE 1 | List of the currently identified TBARs (transcription boundary-associated RNAs).

| TBAR species | Discovery | Length | Reference |
|--------------|-----------|--------|-----------|
| PALRs        | H. sapiens and S. cerevisiae | 200–500 nt | Davis and Ares, 2006; Han et al., 2007 |
| PROMPTs      | H. sapiens | Hundreds of nt | Preker et al., 2008; Preker et al., 2011 |
| PaRNAs       | M. musculus | 40–1,100 nt | Flynn et al., 2011 |
| UaRNAs       | S. cerevisiae | 200–600 nt | Wyers et al., 2005; Neil et al., 2009; Xu et al., 2009 |
| SUTs and CUTs| A. thaliana | – | Chekanova et al., 2007 |
| UNTs         | H. sapiens and M. musculus | – | Kapranov et al., 2007; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009 |
| PASRs        | H. sapiens, M. musculus and A. thaliana | 19–200 nt | Kapranov et al., 2007; Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009; Wang et al., 2011; Ma et al., 2017 |
| TSSaRNAs     | Eukaryotes, bacteria, and archaea | 16–146 nt | Seila et al., 2008; Yus et al., 2012; Park et al., 2014; Zaramela et al., 2014 |
| TiRNAs       | H. sapiens, M. musculus, G. gallus, and D. melanogaster | 13–28 nt | Taft et al., 2009a |
| TARs         | H. sapiens | ~3,100 nt | Yue et al., 2010 |
| TASRs        | H. sapiens, M. musculus, S. pombe, and A. thaliana | 19–200 nt | Kapranov et al., 2007, 2010; Taft et al., 2009a; Leng et al., 2014; Schwab et al., 2016; Ma et al., 2017 |
| TTSaRNAs     | H. sapiens | 22–24 nt | Valen et al., 2011 |

TBARs, transcription boundary-associated RNAs; PARs, promoter-associated RNAs; TARs, terminus-associated RNAs; PALRs, promoter-associated long RNAs; PASRs, promoter-associated small RNAs; TALRs, terminus-associated long RNAs; TASRs, terminus-associated small RNAs; PaRNAs, promoter-associated RNAs; PROMPTs, promoter upstream transcripts; UaRNAs, upstream antisense RNAs; SUTs, stable unannotated transcripts; CUTs, cryptic unstable transcripts; UNTs, upstream non-coding transcripts; TSSaRNAs, transcription start site-associated RNAs; TiRNAs, transcription initiation RNAs; TTSaRNAs, transcription termination site-associated RNAs.
shown that 3’ poly(A) signals are more abundant in regions upstream of the promoter than in those downstream of the promoter. These functional poly(A) signals are involved in the rapid degradation of PROMPTs transcribed in the antisense direction upstream of the associated genes. Thus, the rapid decay of upstream PROMPTs enables efficient elongation of downstream transcripts, which enforces the promoter orientation of protein-coding genes (Ntini et al., 2013). Moreover, if the synthesis of PROMPTs is stalled very early within the TSS-proximal regions, small TSSaRNAs are produced. Small TSSaRNAs are included in the category of PASRs and are discussed below.

**PASRs in Animals**

**Discovery of PASRs**

Promoter-associated small RNAs (sRNAs) identified in human and mouse have been previously named as PASRs (Calabrese et al., 2007; Kapranov et al., 2007). These sRNAs are 20–200 nt in length and originate from the region within −400 bp to +400 bp of the TSS. TSSaRNAs are a type of PASRs that have been identified in a broad spectrum of organisms, including eukaryotes (Seila et al., 2008; Park et al., 2014), bacteria (Yus et al., 2012), and archaea (Zaramela et al., 2014). The length of TSSaRNAs differs between species. For example, in the archaea *Halobacterium salinarum*, the length of TSSaRNAs ranges from 16 to 146 nt, with a median size of 27 nt (Zaramela et al., 2014). In bacteria, the length of TSSaRNAs varies within a narrow range; in *Escherichia coli* and *Mycoplasma pneumoniae*, the length of TSSaRNAs varies from 33 to 40 nt and 35 to 55 nt, respectively (Yus et al., 2012). In murine embryonic stem cells, TSSaRNAs vary in size from 20 to 90 nt. These TSSaRNAs are highly enriched in the region spanning from −250 bp to +50 bp of the TSS (Seila et al., 2008). Thus, sequence features of mouse TSSaRNAs are similar to those described above. Indeed, TSSaRNAs were occasionally confused with the PASRs (Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project, 2009). Another class of PASRs comprises tiRNAs; these are highly conserved in higher metazoans, such as human, chicken, and fruit fly. Similar to the above two PASR classes, tiRNAs are enriched within the region spanning from −60 bp to +120 bp relative to the TSS. However, additional features specific to the tiRNAs have been reported, including their size distribution and GC content (Taft et al., 2009a). The size of tiRNAs varies from 13 to 28 nt, and a significant proportion of these are 18nt in length. Additionally, tiRNAs are GC-rich and show a strong strand bias toward the neighboring TSSs for their biogenesis.

**Biogenesis of PASRs**

The biogenesis of the three types of PASRs described above is Dicer-independent (Calabrese et al., 2007; Seila et al., 2008; Taft et al., 2009a), indicating that the non-canonical sRNA processing pathway(s) might be responsible for their production. Two models have been proposed for the biogenesis of PASRs (Lenhard et al., 2012). The first model is called “backtracking and excision” (Shaevitz et al., 2003; Seila et al., 2009; Taft et al., 2009a, b, 2010; Nechaev et al., 2010). In this model, the elongating RNA Pol II is stalled, inducing it to backtrack toward the upstream TSS after encountering the downstream nucleosome, resulting in a nascent transcript with a short 3’ exposed region. Subsequently, the exposed region is cleaved by the transcription elongation factor SII, resulting in PASR biogenesis. However, in vitro experiments have shown that the pausing and backtracking of Pol II is used to generate 6–14 nt fragments (Lenhard et al., 2012); however, this length is much shorter than that of PASRs. Therefore, further investigation is needed to identify a plausible model. The second model proposed for PASR biogenesis is the “Pol II pausing” model, which is also described as “unsuccessful Pol II elongation followed by RNA degradation” (Buckley et al., 2014; Jonkers et al., 2014). In this model, the RNA elongation complex is stalled at the initiation stage. Pol II pausing results in a nascent transcript without the 5’ cap, thus rendering it susceptible to rapid decay starting from its 5’ end. Only a short RNA sequence physically covered by Pol II is protected from degradation (Valen et al., 2011). The size of this short RNA sequence varies from 17 to 22 nt, which fits well within the median size of PASRs. Notably, the above two models are probably mutually non-exclusive, since different PASR classes may use distinct pathways for their biogenesis.

**Functions of PASRs**

Recent studies have revealed valuable functional implications of PASRs. Depending on the association with certain long non-coding RNAs (lncRNAs) and protein factors, some of the PASRs target specific promoter regions of genes for epigenetic modifications that affect gene transcription (Hamazaki et al., 2017). However, altered transcriptional activity of target genes, in turn, causes fluctuations in PASR biogenesis. Thus, a feedback regulatory loop is established between PASRs and their target genes (Yan and Ma, 2012).

TSSaRNAs have been shown to play an important role in the tissue-specific epigenetic regulation of transcription initiation of target genes (Seila et al., 2008; Taft et al., 2009a, 2011; Henriques et al., 2013). In bacteria, TSSaRNAs are involved in the assembly of functional complexes that enable full-length transcription of their target genes, thus avoiding immature transcription initiation (Yus et al., 2012). The tiRNAs are produced from the promoter regions of highly expressed genes. These sRNAs are associated with chromatin marks representative of active transcription (Taft et al., 2009a, 2011). Specifically, approximately 96% of tiRNAs discovered in human embryonic stem cells overlap with histone H3K4 methylation marks (Taft et al., 2009a). Notably, some of these chromatin marks are generated via the tiRNA-dependent pathway and play important biological roles. For example, in both human and mouse, some of the tiRNAs induce local epigenetic modifications that modulate the localization of TFs that bind to CCCTC motifs (Taft et al., 2011).

**TARS IN ANIMALS: DISCOVERY, BIOGENESIS, AND FUNCTIONS**

In addition to the PARs, transcription signals have also been detected at the 3’ ends of protein-coding genes (Wei et al., 2011) (Table 1). In contrast to PARs, reports on TARS, especially...
terminus-associated long RNAs (TALRs), are limited. In one case of TALRs, Yue et al. (2010) discovered several sense ncRNAs of ~3,100 nt at the 3′ end of a progesterone receptor gene. In another study using human K562 cells, Schwalb et al. (2016) detected transient RNAs downstream of the poly(A) sites using transient transcriptome sequencing (TT-Seq), although the length of these transient RNAs was not described in detail.

The terminus-associated small RNAs (TASRs) ranging from 22 to 200 nt in length are highly conserved in human and mouse (Kapranov et al., 2007, 2010). Traces of sRNAs at the 3′ ends of many animal genes have been reported (Taft et al., 2009a). In fission yeast, the sense TASR snr49 has been detected in the 3′ region of the ribosomal protein-coding gene RPL26 (Leng et al., 2014). Another type of TASRs has been referred to as transcription termination site-associated RNAs (TTSaRNAs) (Valen et al., 2011). The enrichment of AGO-associated sRNAs within the 3′ untranslated regions (UTRs) of protein-coding genes has been observed in human (Valen et al., 2011); these TTSaRNAs typically vary in size from 22 to 24 nt.

Like PASRs, the genomic distribution of TASRs does not indicate a strand-specific bias. Some of the TASRs with non- genomically encoded 5′ poly(U) tails have been reported in human cells (Kapranov et al., 2010). Because these TASRs are antisense to poly(A)-tailed transcripts and are close to the transcription termination sites, these are suggested to be synthesized via an as-yet-unidentified RNA-dependent RNA polymerase (RDR)-dependent pathway. Another mechanism of TASR biogenesis has been reported for the 3′ sense terminus-associated ncRNA snr49 (Leng et al., 2014). Results of scanning deletion analysis within the promoter region reveal that the transcription of snr49 is dependent on the promoter activity of the upstream gene RPL26. Similarly, the transcription of another 3′ sense terminus-associated ncRNA, snr393, relies on the promoter activity of the upstream RPL29 gene. These data suggest that this kind of transcriptional regulation is a conserved mechanism underlying TASR biogenesis (Leng et al., 2014).

To investigate the regulatory roles of TARs overlapping the 3′ termini of protein-coding genes in the human genome, Younger and Corey (2011) designed small duplex RNAs called “microRNA (miRNA) mimics,” which were perfectly complementary to the TARs. The targeting of TARs by “miRNA mimics” affects upstream gene transcription in a dose-dependent manner, highlighting a previously unappreciated role of gene termini and their associated TARs in transcriptional regulation (Younger and Corey, 2011). On the other hand, artificial constructs expressing 3′ UTRs containing miRNA-binding sites act as miRNA sponges (Thomson and Dinger, 2016), indicating that some of the TALRs might function as endogenous miRNA sponges. Another widely accepted model called “gene looping” elucidates TAR-mediated interaction between the 5′ promoters and 3′ terminators. This model provides an unconventional view of transcriptional regulation on a long-range scale (O’Sullivan et al., 2004; Tan-Wong et al., 2008; Tiwari et al., 2008; Yue et al., 2010). The long-range “gene looping” brings the two ends of a gene in close proximity. The sRNAs complementary to the 3′ sense TALRs recruit AGO2 and other protein factors to the 5′–3′ interaction region of the gene, thus modulating its transcription initiation (Yue et al., 2010). This “gene looping” also provides a long-range scaffold, thus enabling communication between endogenous TASRs and promoters of upstream genes (Younger and Corey, 2011).

**Biogenesis of Plant TBARs**

In Arabidopsis, PASRs are located either in upstream or downstream regions of the TSSs (Wang et al., 2011; Ma et al., 2017). In animals, divergent transcription (Beck and Warren, 1988; Core et al., 2008; He et al., 2008; Seila et al., 2008, 2009; Flynn et al., 2011) and bidirectional promoters (Morris et al., 2008; Xu et al., 2009; Wei et al., 2011; Usaka et al., 2014) have been proposed as the mechanisms of PASR biogenesis. Whether these models explain the biogenesis of PASRs in plants needs further investigation. There is no clear model for the biogenesis of
FIGURE 1 | Schematic representation of a uniform annotation system for transcription boundary-associated RNAs (TBARs). In this system, based on their genomic positions, TBARs are first divided into promoter-associated RNA (PARs) and terminus-associated RNA (TARs). Then, according to their sequence length, PARs and TARs are subdivided into long (PALRs and TALRs) and short (PASRs and TASRs) species. For further classification, other important features, such as their biogenesis pathways, modes of action, and biological functions, may be taken into consideration. TBARs, transcription boundary-associated RNAs; PARs, promoter-associated RNAs; TARs, terminus-associated RNAs; PALRs, promoter-associated long RNAs; PASRs, promoter-associated small RNAs; TALRs, terminus-associated long RNAs; TASRs, terminus-associated small RNAs; paRNAs, promoter-associated RNAs; PROMPTs, promoter upstream transcripts; uaRNAs, upstream antisense RNAs; SUTs, stable unannotated transcripts; CUTs, cryptic unstable transcripts; UNTs, upstream non-coding transcripts; TSSaRNAs, transcription start site–associated RNAs; tiRNAs, transcription initiation RNAs; TTSaRNAs, transcription termination site–associated RNAs.

TASRs in animals. It has been suggested that some of the TASRs, especially those that are antisense, are generated through a specific RDR-dependent pathway. Our study in Arabidopsis shows that the accumulation of some TBARs, including PASRs and TASRs, is highly dependent on the activities of RNA Pol IV, RDR2/6, and DCL2/3/4 (Ma et al., 2017).

According to the distribution patterns of TSSs in vertebrates, gene promoters are classified into two categories: sharp and broad. A sharp promoter always has a single predominant TSS, whereas a broad promoter usually has an array of almost equivalent TSSs (Valen et al., 2011; Danino et al., 2015). Interestingly, sharp promoters generate fairly narrow PASR peaks, whereas broad promoters generate PASRs with a much wider distribution range. Notably, in Arabidopsis, both sharp and broad PASR peaks are observed on different gene promoters (Ma et al., 2017). Thus, investigating the relationship between the promoter type and PASR peak shape in plants will be useful.

Functions of Plant TBARs
Few studies have been conducted on the regulatory roles of TBARs in plants. However, studies in animals provide valuable hints on the functions of TBARs in plants. Using the human PR gene as a model, Core et al. (2008) have shown that small duplex RNAs complementary to the ncRNAs originating from the promoter or terminus of the PR gene can efficiently regulate gene transcription. Theoretically, manually designed sRNAs recruit the AGO2 protein to the complementary target ncRNAs (i.e., PALRs or TALRs), which alters the transcription status of the PR gene via the “gene looping” mechanism (Janowski et al., 2005, 2006, 2007; Schwartz et al., 2008; Younger and Corey, 2009; Yue et al., 2010). This regulatory mechanism is conserved in animals (Janowski et al., 2006; Kim et al., 2006; Han et al., 2007; Schwartz et al., 2008; Napoli et al., 2009; Chu et al., 2010). Two research areas need further investigation: first, whether the plant endogenous PASRs and TASRs function in a similar manner as the manually designed sRNAs reported in animals; and second, whether PALRs and TALRs are recognized as targets of PASRs and TASRs, respectively, in plants. The involvement of TBARs in site-specific DNA methylation has been observed in both plants and animals (Mette et al., 2000; Hawkins et al., 2009; Swiezewski et al., 2009; Zheng et al., 2013; Ma et al., 2017), indicating that it is an important pathway for TBAR-mediated regulation of gene transcription.

CONCLUDING REMARKS AND PERSPECTIVES
Need for a Uniform Annotation System for TBAR Research
Owing to the heterogeneity of the TBAR population, many aspects of TBARs including sequence characteristics,
biogenesis pathways and biological functions are poorly understood. One of the most basic and pressing issues derives from their classification. Some of the TBARs such as PASRs, TSSa-RNAs, and tiRNAs exhibit similar genomic locations and sizes in animals. It is unclear whether these sRNAs belong to distinct classes or to the same class (Jacquier, 2009). It is also unclear whether their biogenesis pathways and action modes overlap. Although some of the TBARs are likely to represent “transcription noise,” increasing cases are being reported that emphasize the non-negligible roles of certain TBARs in gene expression regulation. Hence, a uniform annotation system is required to facilitate in-depth research on TBARs. Some criteria have been proposed previously to determine the classification of PARs, including size ranges and positions relative to the TSSs (Kaikkonen et al., 2011; Lenhard et al., 2012). Here, we propose a preliminary system for TBAR annotation (Figure 1).

In this system, we first classified TBARs, based on their genomic position, as PARs and TARs. Then, based on their sequence length, both PARs and TARs were divided into long (>200 nt) and short (<200 nt) species. Using this annotation system, most of the recently reported TBARs could be assigned to one of the categories. For further classification, we suggest that additional important features of TBARs, such as biogenesis pathways, modes of action, and biological functions, should be taken into consideration.

**Future Challenges in TBAR Research**

To advance our understanding of TBARs, further research is needed in a few areas described below. First, it has been suggested that sRNAs mapped to the transcription boundaries of protein-coding genes are degraded remnants of mRNAs (Schwalb et al., 2016). However, considering the non-random enrichment of TBARs with specific sequence length, such as the animal tiRNAs (18 nt) (Taft et al., 2009a) and plant PASRs and TASRs (23–24 nt) (Ma et al., 2017), at least some of the TBARs are unlikely to be generated through random mRNA decay. Nonetheless, a large proportion of TBARs are subjected to rapid degradation after their maturation. According to previous studies (Davis and Ares, 2006; Chekanova et al., 2007; Preker et al., 2008), mutants of RNA decay pathways (e.g., exosome-depleted cell lines) represent promising options for efficient TBAR cloning.

Second, recent evidence shows that a genetic mutation within a protein-coding locus may affect not only the gene of interest but also the associated ncRNAs. Notably, some of these affected ncRNAs, in turn, alter the chromatin state through long-range interactions (Tufarelli et al., 2003; Wei et al., 2011). Thus, when investigating the biological consequences of a mutation within a specific genomic locus, it is necessary to consider the effects on the associated ncRNAs.

Third, as mentioned above, the high frequency of poly(A) signals on PROMPTs is related to their rapid degradation, which has been proposed as one of the mechanisms to ensure unidirectional transcription elongation (Ntini et al., 2013). Since most of the mammalian promoters are bidirectional (Core et al., 2008; Seila et al., 2008; Sigova et al., 2013), functional studies on the involvement of PARs in transcription determination are necessary. Till date, three key factors have been proposed as important players in transcription determination (Wei et al., 2011), including the nucleotide composition, chromatin modifications of promoters, and the “gene looping” mechanism. The nucleotide composition within the promoter region affects the directionality of the promoter (Engstrom et al., 2006). In yeast, the TATA element imposes a constraint on the direction of transcription initiation (Park et al., 2014). Additionally, chromatin modifications induced by the promoter-associated transcription of the ncRNAs may serve as coders for the orientation of transcription elongation. Moreover, the “gene looping” model suggests that long-range interactions link the promoter to its favored 3’ end, thus determining the orientation of transcription elongation (Miele and Dekker, 2008; Laine et al., 2009; Tan-Wong et al., 2009). However, detailed mechanisms of unidirectional transcription elongation require further investigation.

Fourth, in addition to the TBARs associated with protein-coding genes, whether TBARs exist on lncRNAs needs to be revisited. Conserved secondary structures are detected at the ends of certain IncRNAs (Ponting et al., 2009; Yu et al., 2017); however, whether these local structures serve as precursors of PASRs or TASRs remains elusive.

Overall, investigations into all of the above-mentioned challenging but intriguing research areas are needed. Results of these investigations are expected to increase our knowledge of TBARs.

**AUTHOR CONTRIBUTIONS**

DY, XM, HW, and YM: wrote the manuscript. ZZ, DY, and XM: prepared the figure and table. All authors read and approved the final manuscript.

**FUNDING**

This work was funded by the National Natural Science Foundation of China (31571349, 31601062, and 31771436), Zhejiang Provincial Natural Science Foundation of China (LY15C060006) and (LQ16C060003). Hangzhou Scientific and Technological Program (20170432B04).

**REFERENCES**

Affymetrix Encode Transcriptome Project and Cold Spring Harbor Laboratory Encode Transcriptome Project (2009). Post-transcriptional processing generates a diversity of 5’-modified long and short RNAs. *Nature* 457, 1028–1032. doi: 10.1038/nature07759

Amaral, P. P., Dinger, M. E., Mercer, T. R., and Mattick, J. S. (2008). The eukaryotic genome as an RNA machine. *Science* 319, 1787–1789. doi: 10.1126/science.1155472
Amaral, P. P., and Mattick, J. S. (2008). Noncoding RNA in development. *Mamm. Genome* 19, 454–492. doi: 10.1007/s00335-008-9136-7

Andersen, P. K., Jensen, T. H., and Lykke-Andersen, S. (2013). Making ends meet: coordination between RNA 3′-end processing and transcription initiation. *Wiley Interdiscip. Rev. RNA* 4, 233–246. doi: 10.1002/wrna.1156

Beck, C. F., and Warren, R. A. (1988). Divergent promoters, a common form of gene organization. *Microbiol. Rev.* 52, 318–326.

Brown, J. W., Marshall, D. F., and Echeverria, M. (2008). Intronic noncoding RNAs and splicing. *Trends Plant Sci.* 13, 335–342. doi: 10.1016/j.tplants.2008.04.010

Buckley, M. S., Kwak, H., Zipfel, W. R., and Lis, J. T. (2014). Kinetics of promoter Pol II on Hsp70 reveal reveal pausing and key insights into its regulation. *Genes Dev.* 28, 14–19. doi: 10.1101/gad.21886.113

Calabrese, J. M., Sella, A. C., Yeo, G. W., and Sharp, P. A. (2007). RNA sequence analysis defines Dicer’s role in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18097–18102. doi: 10.1073/pnas.0709193104

Chekanova, J. A., Gregory, B. D., Reverdattò, S. V., Chen, H., Kumar, R., Hooker, T., et al. (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* 131, 1340–1353. doi: 10.1016/j.cell.2007.10.056

Chu, Y., Yue, X., Younger, S. T., Janowski, B. A., and Corey, D. R. (2010). Complex Loci in human and mouse genomes. *Cell Cycle* 23, 2604–2609. doi: 10.1016/j.cell.2013.04.028

Chu, Y., Yue, X., Younger, S. T., Janowski, B. A., and Corey, D. R. (2013). Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals. *Mol. Cell.* 52, 517–528. doi: 10.1016/j.molcel.2013.10.001

Huang, V., Qin, Y., Wang, J., Wang, X., Place, R. F., Lin, G., et al. (2010). RNAa is conserved in mammalian cells. *PLoS One* 5:e8848. doi: 10.1371/journal.pone.0008848

Jacquier, A. (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat. Rev. Genet.* 10, 833–844. doi: 10.1038/ nrg2683

Janowski, B. A., Huffman, K. E., Schwartz, J. C., Ram, R., Hardy, D., Shames, D. S., et al. (2005). Inhibiting gene expression at transcription start sites in chromosomal DNA with antigenic RNAs. *Nat. Chem. Biol.* 1, 216–222. doi: 10.1038/nchembio725

Janowski, B. A., Huffman, K. E., Schwartz, J. C., Ram, R., Nordsell, R., Shames, D. S., et al. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat. Struct. Mol. Biol.* 13, 787–792. doi: 10.1038/nsmb1140

Jonkers, I., Kwak, H., and Lis, J. T. (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* 3:e02407. doi: 10.7554/eLife.02407

Kaikkonen, M. U., Lam, M. T., and Glass, K. C. (2011). Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc. Res.* 90, 430–440. doi: 10.1093/cvr/cv097

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484–1488. doi: 10.1126/science.1138341

Kapranov, P., Oszoszl, F., Kim, S. W., Foissac, S., Lipson, D., Hart, C., et al. (2010). New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. *Nature* 466, 642–646. doi: 10.1038/nature09190

Kasab, M. A., Mudassir, M., Singh, A., Muthuraman, N., Bhagat, M., Palanichamy, J. K., et al. (2015). Gene silencing and activation of human papillomavirus 18 is modulated by sense promoter associated RNA in Bidirectionally transcribed long control region. *PLoS One* 10:e0128416. doi: 10.1371/journal.pone.0128416

Kim, D. H., Villeneuve, L. M., Morris, K. V., and Rossi, J. J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat. Struct. Mol. Biol.* 13, 793–797. doi: 10.1038/nsmb1142

Laine, J. P., Singh, B. N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. *Genes Dev.* 23, 2604–2609. doi: 10.1101/gad.1823609

Leng, X. M., Diao, L. T., Li, B., Yi, Z. Y., Chen, C. J., Zhou, H., et al. (2014). The ribosomal protein rpl26 promoter is required for its 3′ sense terminus ncRNA transcription in *Schizosaccharomyces pombe*, implicating a new transcriptional mechanism for ncRNAs. *Biochem. Biophys. Res. Commun.* 444, 86–91. doi: 10.1016/j.bbrc.2014.01.018

Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* 13, 233–245. doi: 10.1038/nrg3163

Li, B., Carey, M., and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* 128, 707–719. doi: 10.1016/j.cell.2007.01.015

Li, L. C., Okino, S. T., Zhao, H., Pookot, D., Place, R. F., Urakami, S., et al. (2006). Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17337–17342. doi: 10.1073/pnas.06070 15103

Li, W., Notani, D., and Rosenfeld, M. G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* 17, 207–223. doi: 10.1038/nrg2016.4

Lykke-Andersen, S., Møbenedo, K. C., and Jensen, T. H. (2011). An ending is a new beginning: transcription termination supports re-initiation. *Cell Cycle* 10, 863–865. doi: 10.4161/cc.10.14931

Ma, X., Han, N., Shao, C., and Meng, Y. (2017). Transcription-wide discovery of PASRs (promoter-associated small RNAs) and TASRs (terminus-associated small RNAs) in *Arabidopsis thaliana*. *PLoS One* 12:e0169212. doi: 10.1371/journal.pone.0169212

Meng, Y., and Shao, C. (2012). Large-scale identification of mirtrons in *Arabidopsis* and rice. *PLoS One* 7:e31163. doi: 10.1371/journal.pone.0031163

Yu et al.
Meng, Y., Shao, C., Wang, H., and Chen, M. (2011). The regulatory activities of plant microRNAs: a more dynamic perspective. *Plant Physiol.* 157, 1583–1595. doi: 10.1104/pp.111.187088

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, A. J. (2006). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* 19, 5194–5201. doi: 10.1038/emboj/2019.19.5194

Miele, A., and Dekker, J. (2008). Long-range chromosomal interactions and gene regulation. *Mol. Biosyst.* 4, 1046–1057. doi: 10.1039/b803580f

Morris, K. V., Santos, S., Turner, A. M., Pastori, C., and Hawkins, P. G. (2008). Bidirectional transcription directly both transcriptional gene activation and suppression in human cells. *PLoS Genet.* 4:e1000258. doi: 10.1371/journal.pgen.1000258

Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., et al. (2006). Breaking barriers to transcription elongation. *Biochim. Biophys. Acta* 1809, 34–45. doi: 10.1016/j.bbagrm.2010.11.001

Napoli, S., Pastori, C., Magistri, M., Carbone, G. M., and Catapano, C. V. (2009). Promoter-specific transcriptional interference and c-myc gene silencing by siRNAs in human cells. *EMBO J.* 28, 1708–1719. doi: 10.1038/embob.2009.139

Nechaev, S., and Adelman, K. (2011). Pol II waiting in the starting gates: regulating the transition from transcription initiation into productive elongation. *Biochim. Biophys. Acta* 1809, 34–45. doi: 10.1016/j.bbagrm.2010.11.001

Ng, A., Core, L. J., and Lis, J. T. (2006). The new feature of active promoters. *Cell Cycle* 5, 2557–2564. doi: 10.4161/cc.5.16.4741

Ponsaerts, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M. S., Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, M. (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2867–2881. doi: 10.1073/pnas.1219940110

Saied, S., Khan, N., Chen, J., Khan, F., and Nasrallah, Y. (2009). Widely bidirectional promoters are the major source of long antisense transcripts of an Arabidopsis Polycomb target. *Nature* 462, 799–802. doi: 10.1038/nature08618

Seila, A. C., Calabrese, J. M., Levine, S. S., Yeo, G. W., Rahl, P. B., Flynn, R. A., et al. (2008). Divergent transcription from active promoters. *Science* 322, 1849–1851. doi: 10.1126/science.1162253

Seila, A. C., Core, L. J., Lis, J. T., and Sharp, P. A. (2009). Divergent transcription: a new feature of active promoters. *Cell Cycle* 8, 2557–2564. doi: 10.4161/cc.8.16.9395

Shaevitz, J. W., Abdouzani, E. A., Landick, R., and Block, S. M. (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature* 426, 684–687. doi: 10.1038/nature02191

Sieg, A. A., Mullin, A. C., Molinie, B., Gupta, S., Orlando, D. A., Guenther, M. G., et al. (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2867–2881. doi: 10.1073/pnas.1219940110

Tafir, R. J., Glazov, E. A., Cloonan, N., Simons, C., Stephen, S., Faulkner, G. J., et al. (2009). Tiny RNAs associated with transcription start sites in animals. *Nat. Genet.* 41, 572–578. doi: 10.1038/ng.312

Tafir, R. J., Kaplan, C. D., Simons, C., and Mattick, J. S. (2009b). Evolution, biogenesis and function of promoter-associated RNAs. *Cell Cycle* 8, 2332–2338. doi: 10.4161/cc.8.15.9154

Tan-Wong, S. M., French, J. D., Proudfoot, N. J., and Brown, M. A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Nat. Protoc.* 3, 5160–5165. doi: 10.1038/nprot.20080104105

Taft, R. J., Hawkins, P. G., Mattick, J. S., and Morris, K. V. (2011). The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. *Epigenetics Chromatin* 4:13. doi: 10.1186/1756-8935-4-13

Taft, R. J., Simons, C., Nakhuri, S., Oey, H., Korbe, D. J., Mercer, T. R., et al. (2010). Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nat. Struct. Mol. Biol.* 17, 1030–1034. doi: 10.1038/nsmb.1841

Tarr, D., and Dinger, M. E. (2016). Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* 17, 272–283. doi: 10.1038/nrg.2016.20

Tiwari, V. K., Mcgarvey, K. M., Lichessi, J. D., Ohm, J. E., Herman, J. G., Schubeler, D., et al. (2008). Pcg proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol.* 6:e206. doi: 10.1371/journal.pbio.0060206

Tufarelli, C., Stanley, J. A., Garrick, D., Sharpe, J. A., Ayuhh, H., Wood, W. G., et al. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat. Genet.* 34, 157–165. doi: 10.1038/ng.1157

Turunen, M. P., Lehtola, T., Heinonen, S. E., Assefa, G. S., Korpisalo, P., Girnary, R., et al. (2009). Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenotherapy. *Circ. Res.* 105, 604–609. doi: 10.1161/CIRCRESAHA.109.200774

Uesaka, M., Agata, K., Oishi, T., Nakashima, K., and Imamura, T. (2017). Evolutionary acquisition of promoter-associated non-coding RNA (pancRNA) repertoires diversifies species-dependent gene activation mechanisms in mammals. *BMCMicrobiol.* 18:285. doi: 10.1186/s12864-017-3662-1

Uesaka, M., Nishimura, O., Go, Y., Nakashima, K., Agata, K., and Imamura, T. (2014). Bidirectional promoters are the major source of gene activation-associated non-coding RNAs in mammals. *BMCMicrobiol.* 15:35. doi: 10.1186/1471-2164-15-35

Valen, E., Preker, P., Andersen, P. R., Zhao, X., Chen, Y., Ender, C., et al. (2011). Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. *Nat. Biotechnol.* 29, 1057–1062. doi: 10.1038/nbt.1984

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., et al. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126–130. doi: 10.1038/nature06992

Yu et al.
Wang, X., Laurie, J. D., Liu, T., Wentz, J., and Liu, X. S. (2011). Computational dissection of Arabidopsis smRNAome leads to discovery of novel microRNAs and short interfering RNAs associated with transcription start sites. *Genomics* 97, 235–243. doi: 10.1016/j.ygeno.2011.01.006

Wei, W., Pelechano, V., Jarvelin, A. I., and Steinmetz, L. M. (2011). Functional consequences of bidirectional promoters. *Trends Genet.* 27, 267–276. doi: 10.1016/j.tig.2011.04.002

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., et al. (2005). Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* 121, 725–737. doi: 10.1016/j.cell.2005.04.030

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., et al. (2009). Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457, 1033–1037. doi: 10.1038/nature07728

Yan, B. X., and Ma, J. X. (2012). Promoter-associated RNAs and promoter-targeted RNAs. *Cell. Mol. Life Sci.* 69, 2833–2842. doi: 10.1007/s00018-012-0953-1

Younger, S. T., and Corey, D. R. (2009). The puzzle of RNAs that target gene promoters. *ChemBioChem* 10, 1135–1139. doi: 10.1002/cbic.20090015

Younger, S. T., and Corey, D. R. (2011). Transcriptional regulation by miRNA mimics that target sequences downstream of gene termini. *Mol. Biosyst.* 7, 2383–2388. doi: 10.1039/c1mb05090g

Yu, D., Ma, X., Zuo, Z., Shao, W., Wang, H., and Meng, Y. (2017). Bioinformatics resources for deciphering the biogenesis and action pathways of plant small RNAs. *Rice* 10:38. doi: 10.1186/s12284-017-0177-y

Yue, X., Schwartz, J. C., Chu, Y., Younger, S. T., Gagnon, K. T., Elbashir, S., et al. (2010). Transcriptional regulation by small RNAs at sequences downstream from 3′ gene termini. *Nat. Chem. Biol.* 6, 621–629. doi: 10.1038/nchembio.400

Yus, E., Guell, M., Vivancos, A. P., Chen, W. H., Lluch-Senar, M., Delgado, J., et al. (2012). Transcription start site associated RNAs in bacteria. *Mol. Syst. Biol.* 8:585. doi: 10.1038/msb.2012.16

Zaramela, L. S., Vencio, R. Z., ten-Caten, F., Baliga, N. S., and Koide, T. (2014). Transcription start site associated RNAs (TSSaRNAs) are ubiquitous in all domains of life. *PLoS One* 9:e107680. doi: 10.1371/journal.pone.0107680

Zeitlinger, J., Stark, A., Kellis, M., Hong, J. W., Nechaev, S., Adelman, K., et al. (2007). RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* 39, 1512–1516. doi: 10.1038/ng.2007.26

Zheng, Q., Rowley, M. J., Bohmdorfer, G., Sandhu, D., Gregory, B. D., and Wierzbicki, A. T. (2013). RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. *Plant J.* 73, 179–189. doi: 10.1111/tpj.12034

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Yu, Ma, Zuo, Wang and Meng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.