



Terminus-Associated Non-coding RNAs: Trash or Treasure?

Wen-Juan Ni, Fuhua Xie* and Xiao-Min Leng*

School of Basic Medicine, Gannan Medical University, Ganzhou, China

OPEN ACCESS

Edited by:

Nikolay Shirokikh,
Australian National University,
Australia

Reviewed by:

Vladimir P. Zhdanov,
Boreskov Institute of Catalysis (RAS),
Russia
Lucy Barrett,
University of Western Australia,
Australia

Takuya Imamura,
Hiroshima University, Japan

*Correspondence:

Fuhua Xie
fuhuax2004@163.com
Xiao-Min Leng
lengxm@mail2.sysu.edu.cn

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 16 April 2020

Accepted: 25 August 2020

Published: 15 September 2020

Citation:

Ni W-J, Xie F and Leng X-M
(2020) Terminus-Associated
Non-coding RNAs: Trash or Treasure?
Front. Genet. 11:552444.
doi: 10.3389/fgene.2020.552444

3' untranslated regions (3' UTRs) of protein-coding genes are well known for their important roles in determining the fate of mRNAs in diverse processes, including trafficking, stabilization, translation, and RNA-protein interactions. However, non-coding RNAs (ncRNAs) scattered around 3' termini of the protein-coding genes, here referred to as terminus-associated non-coding RNAs (TANRs), have not attracted wide attention in RNA research. Indeed, whether TANRs are transcriptional noise, degraded mRNA products, alternative 3' UTRs, or functional molecules has remained unclear for a long time. As a new category of ncRNAs, TANRs are widespread, abundant, and conserved in diverse eukaryotes. The biogenesis of TANRs mainly follows the same promoter model, the RNA-dependent RNA polymerase activity-dependent model, or the independent promoter model. Functional studies of TANRs suggested that they are significantly involved in the versatile regulation of gene expression. For instance, at the transcriptional level, they can lead to transcriptional interference, induce the formation of gene loops, and participate in transcriptional termination. Furthermore, at the posttranscriptional level, they can act as microRNA sponges, and guide cleavage or modification of target RNAs. Here, we review current knowledge of the potential role of TANRs in the modulation of gene expression. In this review, we comprehensively summarize the current state of knowledge about TANRs, and discuss TANR nomenclature, relation to ncRNAs, cross-talk biogenesis pathways and potential functions. We further outline directions of future studies of TANRs, to promote investigations of this emerging and enigmatic category of RNA.

Keywords: 3' termini, 3' UTR, ncRNA, biogenesis, function

INTRODUCTION

The encyclopedia of DNA elements (ENCODE) project aims to reveal functional elements of the human genome, thereby providing new insights into gene and genome functions (Ding et al., 2014; Moraes and Goes, 2016). For instance, RNA sequencing revealed that eukaryotic genomes are pervasively transcribed, using different regions to generate abundant and versatile non-coding RNAs (ncRNAs) (Kapranov et al., 2007; Forrest and Carninci, 2009; Jacquier, 2009; Clark et al., 2011; Jensen et al., 2013; Lu and Lin, 2019). Well-characterized ncRNAs, such as long non-coding RNAs (lncRNAs), small nucleolar RNAs (snoRNAs), and microRNAs (miRNAs), have been found to be variably produced. lncRNAs mainly derive from intergenic regions, introns, and antisense strands (Ayupe et al., 2015; Liu et al., 2015). SnoRNAs mainly arise from introns and

intergenic regions. The possible origins of miRNAs resemble those of snoRNAs. Consistently, similar percentages of intronic snoRNAs and intronic miRNAs have been reported in different eukaryotes (Mattick, 2003; Brown et al., 2008; Scott and Ono, 2011). Additionally, many new ncRNAs located at the 3' and 5' termini of genes have also been detected (Kapranov et al., 2007; Jacquier, 2009; Djebali et al., 2012; Ma et al., 2017; Laudadio et al., 2018). Owing to the absence of specific patterns in most 3' end-associated ncRNAs and the limitations of the RNA sequencing technologies, these ncRNAs have usually been ignored for the past decade.

Investigation of the full landscape of 3' untranslated regions (3' UTRs) across species and cell types has contributed substantially to our understanding of their biogenesis and functions. Studies on the functions of 3' UTRs focused primarily on their role in the regulation of gene expression, including mRNA trafficking, translational control, metabolism, and mRNA-protein structures (Wickens et al., 1997; Andreassi and Riccio, 2009; Denti et al., 2013; Jia et al., 2013; Pánek et al., 2016; Mayr, 2017). However, ncRNAs found around 3' termini are usually not identified as biologically important. Indeed, the presence of terminus-associated small RNAs (TASRs) in both human and mouse genomes was firstly reported in 2007. These RNAs are usually scattered at both strands of protein-coding genes and do not exhibit unique lengths, specific base compositions, or typical secondary structures (Kapranov et al., 2007). Other small RNAs have also been detected at the 3' ends of genes in both human and chicken genomes (Taft et al., 2009; Wei et al., 2011). Interestingly, these 3' end-associated small RNAs are significantly different from the characteristic transcription initiation RNAs (tiRNAs) (Taft et al., 2009). In addition, transient transcriptome sequencing (TT-seq) has detected short-lived RNAs downstream of the polyadenylation [poly(A)] sites in human K562 cells. However, these RNAs are difficult to detect as they are usually cleaved from these sites, resulting in unprotected 5' ends (Schwalb et al., 2016). Thus, these terminus-associated non-coding RNAs (TANRs) did not attract attention due to the lacking of unique length ranges and typical secondary structures.

From a technical perspective, transcriptome sequencing and microarrays show limitations for the discovery of TANRs. Indeed, transcriptome sequencing requires the construction of cDNA libraries and TANRs are often discarded during the rRNA removal step of this process or mixed with annotated transcript fragments afterward. In fact, as mixed fragments, they can partially or completely overlap with the annotated transcripts. Overlapping RNAs can be mapped as part of the annotated transcripts, alternative UTR regions, or even discarded. Furthermore, if some TANRs do not overlap with annotated transcripts, these would be filtered out as erroneous transcripts during bioinformatic analyses. In similarity, transcriptome microarrays are based on available information on annotated transcripts, usually excluding TANRs. Hence, TANRs have been mostly ignored in gene expression studies, given their lack of specific patterns, the uncertainty of their transcriptional origin, and other methodological difficulties (Jacquier, 2009; Yu et al., 2018).

Although TANRs usually are not identified as high-value targets, increasing evidence implied that they are important molecules for several cellular activities. For instance, the detection of diverse TANRs in eukaryotes suggested that they are widespread, abundant, and conserved. Moreover, studies of their biogenesis and functions pointed at TANRs as versatile molecules regulating gene expression. Since the biogenesis and functions of many TANRs are still unclear, and an increasing number of TANRs have been reported, elucidating their biological functions and mechanisms of action has become a new frontier in the field of RNA research.

DISCOVERY OF TANRs

Applications and breakthroughs of next-generation sequencing (NGS) and gene array in transcriptomics have revealed eukaryotic genomes can generate a multitude of diverse RNA species (Willingham et al., 2006; Kapranov et al., 2007; Jacquier, 2009). Owing to the presence of bidirectional promoters, one more lncRNA and many small ncRNAs have been found around the corresponding mRNA transcription start sites (TSSs) (Seila et al., 2008; Neil et al., 2009; Xu et al., 2009) (**Figure 1**). These RNAs can be generally termed as promoter-associated RNAs (PARs), including promoter-associated non-coding RNAs (pancRNAs) (Yamamoto et al., 2016; Uesaka et al., 2017), promoter upstream transcripts (PROMPTs) (Preker et al., 2011), upstream antisense RNAs (UaRNAs) (Flynn et al., 2011), stable unannotated transcripts (SUTs), cryptic unstable transcripts (CUTs) (Neil et al., 2009; Xu et al., 2009), promoter-associated long RNAs (PALRs) (Kapranov et al., 2007), tiRNAs (Taft et al., 2009), and other PARs (Jiang et al., 2007). When the attention was shifted to the 3' terminus, diverse ncRNAs were also discovered. These were divided into different subclasses: TASRs (Kapranov et al., 2007), antisense TASRs (aTASRs) (Kapranov et al., 2010), terminus-associated small nucleolar RNAs (TASNrs) (Leng et al., 2014), transcription termination site associated RNAs (TTSa-RNAs) (Valen et al., 2011; Laudadio et al., 2018), transcription boundary-associated RNAs (TBARs) (Yu et al., 2018), terminus-associated long RNAs (TALRs) (Yue et al., 2010), and 3' UTR-associated RNAs (uaRNAs) (Mercer et al., 2011) (**Figure 2**). For a clearer distinction, we highlight that the abbreviation "UaRNAs" stands for "upstream antisense RNAs," while "uaRNAs" indicates "3' UTR-associated RNAs." The different methods used for TANR identification together with the main characteristics of TANRs are summarized in **Table 1**.

Terminus-associated small RNAs were firstly reported to cluster at the 3' termini of mRNAs (Kapranov et al., 2007). In mammals, there are about 200 TASR copies per cell (total numbers for all protein-coding genes), constituting approximately 3% of the small RNA library (Kapranov et al., 2007; Djebali et al., 2012). In *Arabidopsis thaliana*, TASR peaks were identified on 287 protein-coding genes, demonstrating that TASRs mainly accumulated in leaves and young seedlings (Ma et al., 2017). Altogether, the relevance of TASRs in both mammals and plants has been underestimated and the study of these small

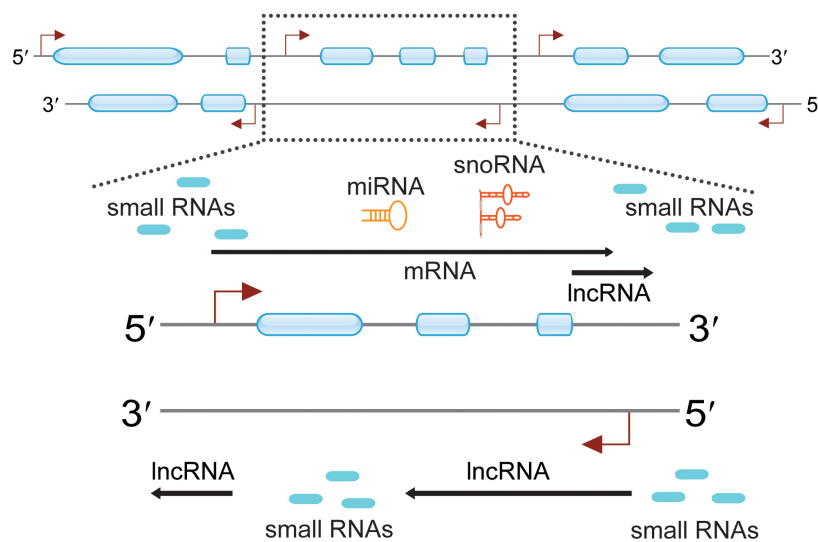


FIGURE 1 | Pervasive transcription across eukaryotic genomes can generate a multitude of diverse RNA species. Genomic regions are indicated with two thin lines marked with direction (5' to 3'). Exons are presented as blue boxes, and transcription start sites (TTS) are indicated with red angled arrows. Certain regions of the genome were highlighted and indicated with dotted frame lines. The zoomed in regions enlist many ncRNAs are observed around the protein-coding genes. These ncRNAs include small RNAs, lncRNAs, miRNAs, and snoRNAs.

RNAs did not receive priority in recent years for their lack of specific patterns. Considering their specific location within 3' UTRs and the presence of poly(A) tails at their 3' ends, TASRs have been regarded as degraded mRNA products or alternative 3' UTRs in many studies.

Interestingly, a novel type of TASRs containing polyU tails at their 5' end have been identified and renamed aTASRs, because they are antisense to 3' UTRs (Kapranov et al., 2010). There are about 600 aTASR copies per cell, corresponding to 702 RefSeq-annotated protein-coding genes. Meanwhile, 1258 transcripts with non-genomically encoded 5' poly(U) stretches closely associated with the 3' termini of known RNAs can also be found in the UCSC Genome Browser database (Kapranov et al., 2010). Since aTASRs display a stretch of U residues at their 5' ends but no poly(A) at their 3' ends, they would be discarded in a conventional transcriptome analysis or library construction. Thus, direct RNA sequencing without prior conversion of RNA to cDNA would facilitate the discovery of novel ncRNAs (Furlan et al., 2020).

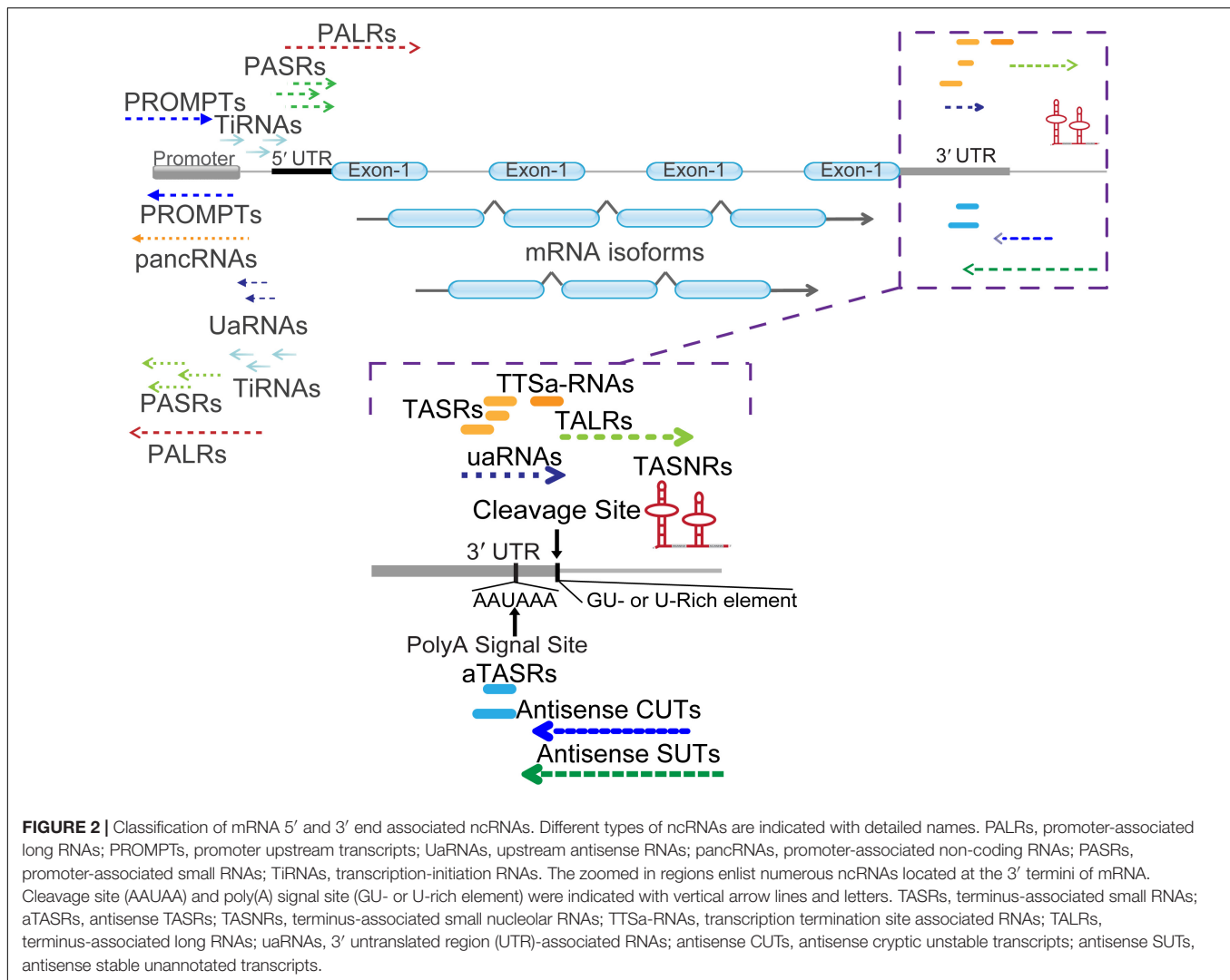
Argonaute (AGO) proteins are highly specialized binding small RNAs and can regulate gene expression at both transcriptional and posttranscriptional level by interacting with other proteins (Meister, 2013). By sequencing AGO1/2 immunoprecipitated libraries, several TTSa-RNAs were identified in *Homo sapiens*, particularly clustered close to the 3' termination sites of mRNAs (Valen et al., 2011). Such TTSa-RNAs were found to be originated from 2822 protein-coding genes on average. Additionally, TTSa-RNAs are rich in G residues at their 5' end and have a peculiar oligo(A) tail at their 3' end (Laudadio et al., 2018). Compared to TASRs and aTASRs, TTSa-RNAs display shorter lengths (22 to 24 nt) and a specific cellular localization (enriched in nucleus). Beyond linear

TASRs, aTASRs, and TTSa-RNAs, hairpin TANRs (for some given genes) have been found in the yeast species related to *Schizosaccharomyces pombe* Lindner (Leng et al., 2014). In addition to small RNAs, lncRNAs, such as TALRs (for a given gene), uaRNAs (3' UTR-associated RNAs) (about 1000 copies per cell on average in human), antisense CUTs, and SUTs (about 1000 copies per cell on average), have also been reported (Neil et al., 2009; Xu et al., 2009; Yue et al., 2010; Mercer et al., 2011). Given the evolutionary pressure toward the conservation of 3' UTR regions, TANRs are usually conserved among different species.

Regarding the genomic location of these ncRNAs, TASRs, aTASRs, TTSa-RNAs, and uaRNAs are located within 3' UTRs. In particular, TANRs and aTASRs start from poly(A) signal sites, while TTSa-RNAs end at the cleavage sites. Furthermore, TALRs and a small subset of antisense CUTs/SUTs usually overlap with 3' UTRs. On the other hand, TANRs are located downstream of 3' UTRs. As indicated by their name, aTASRs and antisense CUTs/SUTs are located on the antisense strand, while other ncRNAs are located on the sense strand (Figure 2). According to their length (more or less than 200 nt), TALRs, uaRNAs, and antisense CUTs/SUTs are classified as lncRNAs, whereas others are considered small RNAs. Overall, TANRs vary considerably in their genomic location, strand, and length (Table 1).

BIOGENESIS OF TANRs

Studies of the biogenesis of ncRNAs are required to elucidate their functions and potential roles in the regulation of gene expression (Kim et al., 2009; Li et al., 2010). MiRNAs are currently the best-described small regulatory ncRNAs that follow a specific biogenesis pathway, requiring DROSHA/DGCR8, DICER1, and



AGO proteins (Daugaard and Hansen, 2017; Saeed et al., 2020). As TANRs are a novel class of ncRNAs, most but not all proteins associated to their biogenesis are unknown. According to their maturation process, the biogenesis of TANRs can generally occur by one of three models: the same promoter model, the RdRP activity-dependent model, and the independent promoter model.

The Same Promoter Model

In the same promoter model of TANR biogenesis, firstly, the transcription of TANR precursors is coupled to that of the upstream mRNAs using the same promoter. Then, maturation of TANRs occurs by posttranscriptional cleavage. Considering that the maturation processes of TASRs, TASNRs, TTSa-RNAs, TALRs, and uaRNAs share many characteristics, we summarize them altogether.

Terminus-associated small RNAs are located within the 3' UTR of genes where no histone modifications marking active promoters or enrichment for RNA polymerase II (RNAPII) occupancy are found (Mercer et al., 2011). Hence, it is reasonable to infer that for their maturation TASRs undergo

posttranscriptional cleavage. Studies on the biogenesis of TALRs and TTSa-RNAs also suggested that their maturation mainly depends on posttranscriptional cleavage from the corresponding mRNAs (Yue et al., 2010; Laudadio et al., 2018). However, this maturation process significantly differs from that of miRNAs. Firstly, evidence of the formation of secondary structures and of the corresponding passenger strands, characteristic of miRNA maturation, has not been found for these ncRNAs (Valen et al., 2011). Furthermore, genome-wide studies of TTSa-RNAs also determined that the regions flanking TTSa-RNAs do not tend to form hairpin structures more than randomly picked genomic regions (Laudadio et al., 2018). Secondly, altered expression of DICER and AGO2, required for miRNA biogenesis, had no effects on TTSa-RNA biogenesis (Valen et al., 2011; Laudadio et al., 2018). Importantly, defined sites within the polyA tail and approximately 75% of mRNA 3' ends carry at least one TTSa-RNA read, suggesting that mRNA 3' end processing is involved in their biogenesis (Valen et al., 2011). However, TTSa-RNAs are not by-products of mRNA degradation, since they display upstream poly(A) signals and are specifically loaded on AGO

TABLE 1 | List of TANRs described in this article.

TANRs	Technology	5' Cap	3' polyA	Species	Length	Strand	Structure	References
TASRs	Tiling array	No	Yes	<i>H. sapiens</i> , <i>M. musculus</i> and <i>A. thaliana</i>	22–200 nt	Sense	Linear	Kapranov et al. (2007); Djebali et al. (2012), Ma et al. (2017)
aTASRs	Helicos single-molecule sequencing	No	No	<i>H. sapiens</i> and <i>M. musculus</i>	<200 nt	Antisense	Linear	Kapranov et al. (2010)
TASNRs	Northern blot, RT-PCR, RACE	No	No	<i>Schizosaccharomyces group</i>	<200 nt	Sense	Hairpin	Leng et al. (2014)
TTSa-RNAs	RNA sequencing	No	No	<i>H. sapiens</i>	22–24 nt	Sense	Linear	Valen et al. (2011); Laudadio et al. (2018)
TALRs	RACE, RT-PCR	Yes	Yes	<i>H. sapiens</i>	>200 nt	Sense	Linear	Yue et al. (2010)
uaRNAs	CAGE, SAGE, Microarray	Yes	Yes	<i>H. sapiens</i> , <i>M. musculus</i> and <i>D. melanogaster</i>	>200 nt	Sense	Linear	Mercer et al. (2011)
Antisense CUTs/SUTs	Tiling array, RNA sequencing	Yes	Yes	<i>S. cerevisiae</i>	200–500 nt	Antisense	Linear	Neil et al. (2009); Xu et al. (2009)

proteins. Moreover, TTSa-RNAs tend to carry a G residue in the first position at the 5' end and an oligo(A) tail (four or more As) at the 3' end, supporting the hypothesis that TTSa-RNAs undergo posttranscriptional cleavage from the corresponding mRNAs (Laudadio et al., 2018).

Notably, detailed studies on the biogenesis of TASNRs and uaRNAs strongly indicated the same promoter model as the typical one for the biogenesis of most TANRs. In particular, two TASNR precursors (*rpl26-snR49* and *rpl29-snR93*) highly overlapped with upstream mRNAs; no promoters were detected between mature TASNRs and their precursors; and promoter deletion analysis confirmed that the precursor of TASNR *snR49* and the corresponding upstream mRNA used the same promoter for the regulation of their transcription. Thus, TASNRs undergo processing from precursors during maturation (Leng et al., 2014). As for uaRNAs, no active promoters or enrichment for RNAPII occupancy have been found within the 3' UTR; however, exon-intron junctions have been detected (Mercer et al., 2011). Moreover, a detailed study on the biogenesis of the uaRNA *FLJ11812* in human cells confirmed that the maturation of this ncRNA depends on posttranscriptional cleavage, and that the TIA1 protein is responsible for this process (Ge et al., 2014).

Although TASNRs and uaRNAs exploit the same promoters of their respective upstream protein-coding genes for transcription, their precursors originate differently. Indeed, uaRNAs may derive from their corresponding mRNAs through maturation by cleavage similarly to TTSa-RNAs. Conversely, TASNR precursors are different transcripts from their corresponding mRNAs, although highly overlapping. As for TASRs and TALRs, it is still unknown whether they are cleaved from their corresponding mRNAs.

The RdRP Activity-Dependent Model

The RNA-dependent RNA polymerase (RdRP) plays a key role in RNA silencing in fungi, plants, and worms by generating double-stranded RNAs (dsRNAs) from RNA templates (Duempelmann et al., 2020). In the RdRP activity-dependent model of TANR biogenesis, RdRP can *de novo* synthesize antisense TANRs at the 3' termini of mRNAs by using the sense mRNAs as templates. For instance, it has been reported that the human telomerase reverse transcriptase (TERT) RdRP can perform *de novo* synthesis

of short interfering RNAs (siRNAs) that are complementary to template RNAs (Maida et al., 2016). Thus, *de novo* RNA synthesis by RdRP suggests the existence of a novel RNA copying mechanism. Recent studies strongly indicated that the biogenesis of aTASRs depends on RdRP. Indeed, aTASRs contain non-genomically encoded poly(U) stretches at their 5' ends that are complementary to the 3' poly(A) tails of mRNAs (Kapranov et al., 2010). These double-stranded and complementary RNAs have been detected in both human cells and plants (Kapranov et al., 2010; Ma et al., 2017). In *A. thaliana*, aTASR fragments were preferentially incorporated into AGO4 and aTASR accumulation was significantly decreased in *rdr2* (RNA-dependent RNA polymerase 2), *nrpd1a* (RNA polymerase IVa), and *nrpd1b* (RNA polymerase IVb) mutants. Thus, RdRPs and RNA polymerase IV are responsible for the biogenesis of some aTASRs, even though the detailed mechanisms remain unknown (Ma et al., 2017). However, the endogenous biochemical pathway that mediates copying of aTASRs in human cells still requires further investigation (Kapranov et al., 2010).

The Independent Promoter Model

In the independent promoter model, TANRs on the antisense strand have their own promoters. As independent transcripts, their biogenesis is usually regulated by their upstream promoter regions. Although TANRs include only a small number of antisense CUTs/SUTs, several studies indicated that independent promoters are primarily responsible for their biogenesis in *Saccharomyces cerevisiae*. This conclusion derived from the fact that the transcriptional initiation sites of antisense CUTs or SUTs are located in nucleosome-free regions (NFRs), corresponding to promoter regions. Thus, independent transcription is the main biogenesis mechanism of antisense CUTs or SUTs (Neil et al., 2009; Xu et al., 2009).

FUNCTIONS OF TANRs

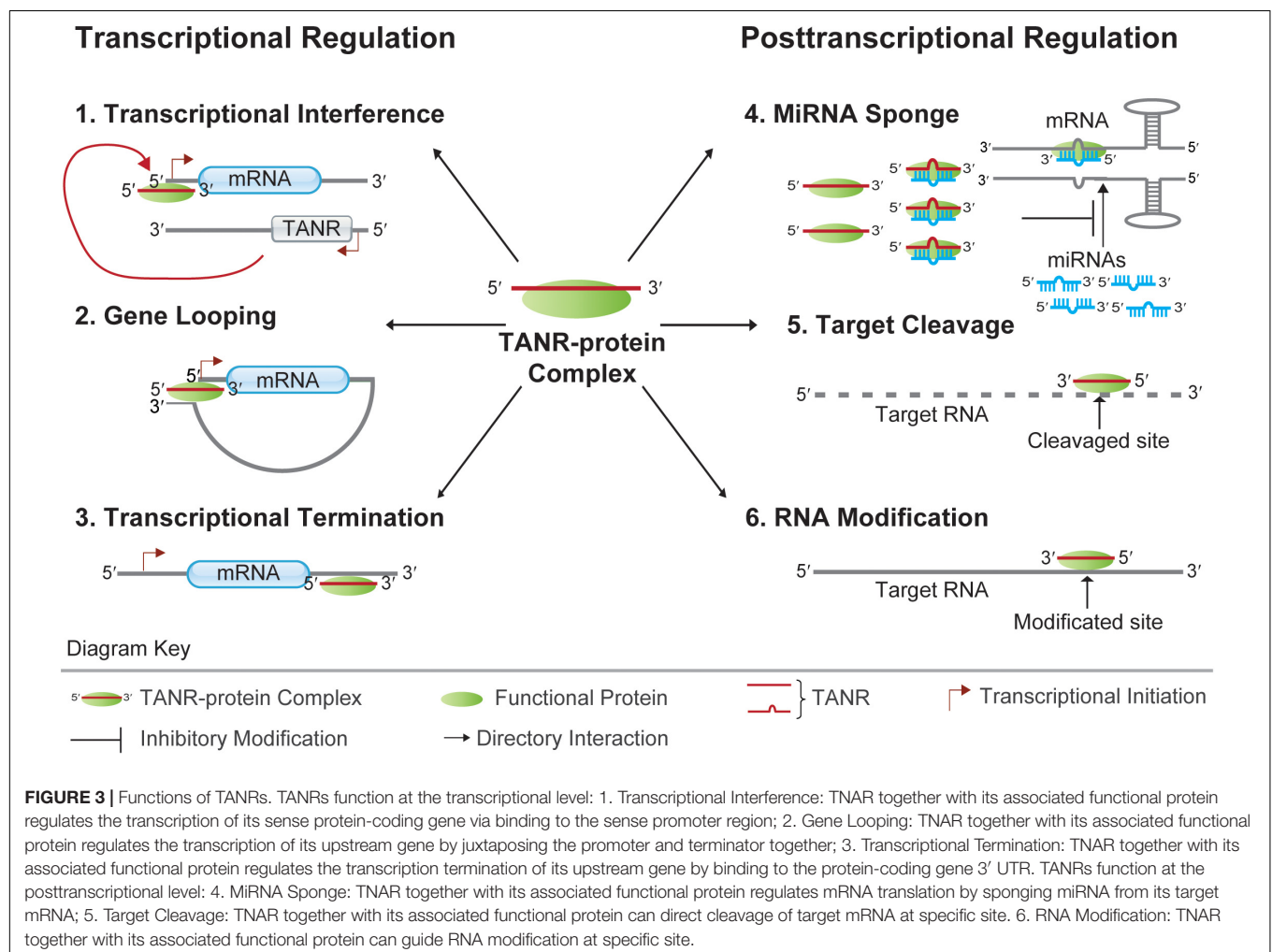
The existence of different pathways of TANR biogenesis suggests that they are important for some cellular activities. NcRNAs typically function by forming various ribonucleoproteins (RNPs) together with several proteins. Well-known functional RNP

particles include snoRNA ribonucleoproteins (snoRNPs) and miRNA-AGO ribonucleoproteins (miRNPs). These RNPs contain the respective RNAs and a small set of associated proteins (Bachellerie et al., 2002; Bartel, 2004). Within miRNPs, miRNAs usually cause degradation and translational repression of target mRNAs through the formation of miRNA-mRNA duplexes. However, miRNA-mRNA interactions are dynamically regulated by different physiological or pathological conditions (Ni and Leng, 2015). As a group of widely studied functional proteins, AGO proteins associate with a diverse variety of ncRNAs, thereby providing functional and regulatory support for ncRNA-mediated modulation of gene expression (Joshua-Tor and Hannon, 2011; Daugaard and Hansen, 2017). It was reported that TANRs enriched in different subcellular compartments (cytoplasm and nucleus) can interact with different AGO proteins in eukaryotes. Hence, TANRs may regulate gene expression at both transcriptional and posttranscriptional levels (Figure 3).

Transcriptional Regulation by TANRs

Members of eukaryotic AGO protein family are key players of gene expression (Meister, 2013). Interestingly, a previous study

showed that synthetic small RNAs fully complementary to a TALR located beyond the 3' terminus of progesterone receptor (*PR*) mRNA could modulate *PR* transcription (Younger and Corey, 2011). This provides new insights into the function of TANRs with high nuclear localization. Firstly, the TALR is loaded onto AGO2 upon addition of exogenous miRNA mimics. Then, the complex formed of miRNA mimics, TALR, and AGO2 is recruited to the promoter region of an upstream gene. Finally, a gene loop juxtaposing the promoter and terminator is formed, resulting in altered regulation of transcription (Figure 3) (Yue et al., 2010). Notably, the formation of gene loops is thought to mediate long-distance transcriptional regulation in different eukaryotes (Bratkowski et al., 2018). However, functional studies of AGO1 and AGO2-associated TTSa-RNAs strongly argued against their specific recruitment on chromatin given their nucleoplasm/chromatin abundance, although Gene Ontology (GO) analysis suggested that genes giving rise to TTSa-RNAs are significantly enriched in the regulation of cell cycle progression and DNA integrity checkpoints (Laudadio et al., 2018). On the other hand, evidence of transcriptional stalling via RNAPII backtracking triggering nucleolytic degradation of the nascent RNA indicates that TTSa-RNAs may be implied in



the termination of mRNA transcription (Valen et al., 2011). For instance, a recent study in *A. thaliana* indicated that promoter-proximal RNAPII stalling can regulate plant gene transcription (Thomas et al., 2020). Thus, it is reasonable to infer that mammalian TTSa-RNAs might participate in the regulation of gene transcription through the modulation of transcriptional termination (Figure 3).

In plants, AGO1 represses target RNAs in the cytoplasm, while AGO4 usually directs *de novo* DNA methylation in the nucleus (Baulcombe, 2004; Vaucheret, 2008; Carbonell, 2017). Site-specific DNA methylation signals were observed on several genomic loci corresponding to the peaks of many TASRs associated with AGO4 in *A. thaliana* (Ma et al., 2017). Furthermore, some aTASRs are preferentially incorporated into AGO4. Thus, a subset of the TASRs and aTASRs reported in *A. thaliana* may be involved in site-specific DNA methylation (Ma et al., 2017). However, it is not clear if TANR-mediated gene looping is required to guide DNA methylation.

In *S. cerevisiae*, antisense CUTs/SUTs usually couple the transcriptional regulation of neighboring genes. As overlapping and divergent transcripts, they may act as local regulatory signals for transcriptional interference (Figure 3) (Neil et al., 2009; Xu et al., 2009). In addition, transcriptional interference mediated by *cis*-acting antisense CUTs/SUTs involves several chromatin modifiers (such as Set2p, Set1p, Rco1p, and Eaf3p) (Nevers et al., 2018). In a recent related report, the transcription of approximately 20% of *S. cerevisiae* genes was found to be repressed by antisense ncRNAs via a chromatin-based transcription interference mechanism. Hence, using near-base-pair-resolution techniques in antisense CUTs/SUTs-inducible strains would reveal the relationship between antisense transcription and repression of sense gene expression, nucleosome occupancy, and transcription-associated histone modifications (Gill et al., 2020).

Posttranscriptional Regulation by TANRs

Previous reports suggested that 3' UTRs can function *in trans* to regulate cell proliferation and differentiation in the absence of corresponding protein-coding transcripts (Rastinejad and Blau, 1993; Amack et al., 1999; Jenny et al., 2006). For example, expression of oskar 3' UTR in *Drosophila* could rescue the eggless defect of oskar null-mutants in the absence of the Oskar protein. Indeed, the oskar 3' UTR functions as a scaffold for trafficking and accumulation of Staufen during oogenesis (Jenny et al., 2006). Moreover, in *A. thaliana*, the overexpressed ncRNA *IPS1* can act as a competing endogenous RNA (ceRNA) that positively regulates the expression of *PHO2* by sequestering *miR-399* from its target site (Franco-Zorrilla et al., 2007). Also, in human embryonic stem cells (hESCs) and human umbilical vein endothelial cells (HUVECs), the uaRNA *FLJ11812* derived from the 3' UTR of *TGFB2* can be targeted by *miR-4459*. Conversely, uaRNA *FLJ11812* can upregulate the levels of the proteins CDC20B and ATG13, whose coding genes can also be targeted by *miR-4459*. Thus, this uaRNA acts as a ceRNA by sponging *miR-4459* from its target mRNAs (Lu et al., 2015). Therefore, uaRNAs can act as decoys to sponge miRNAs from their target mRNAs (Figure 3). Alternatively, they may act as

scaffolds to form regulatory RNA-protein complexes that are functional even in the absence of their corresponding proteins (Mercer et al., 2011).

Structural and functional analyses of ncRNAs in fission yeast suggested that some TANRs act as guide snoRNAs. By forming specific snoRNPs, these snoRNAs can direct methylation or pseudouridylation of target RNAs. Notably, most of such site-specific modifications can affect cell growth *in vivo*. For example, TANR *snR49* was predicted to mediate pseudouridylation of 18S rRNA at the U121 and U305 sites. Upon deletion of TANR *snR49*, the corresponding modifications on rRNA disappeared with consequent delay of cell growth. Furthermore, posttranscriptional modifications of target RNAs by TANRs are conserved in yeasts (Leng et al., 2014). Thus, TANRs can act as guide RNAs for targeted RNA modifications (Figure 3).

Since the production of human aTASRs is positively correlated with that of their associated mRNAs, functional studies of aTASRs were based on the corresponding transcripts. These transcripts corresponded to functionally annotated proteins and were further analyzed. Functional enrichment analysis suggested that they are related to translation. Indeed, the GO categories of "structural constituent of ribosome," "translation," and "RNA binding" were all significantly overrepresented. Due to the bias of enrichment analysis toward highly synthesized transcripts, all human genes were used as background for a second estimation of enrichment. Nevertheless, similar results were obtained, with the GO biological function category "translation" scoring as the top hit (Kapranov et al., 2010). In *A. thaliana*, aTASRs associated with cytoplasmic AGO1 are proposed to mediate target RNA cleavage (Figure 3) (Ma et al., 2017). Perhaps, synthetic aTASR mimics would help to reveal their mechanism of translational regulation.

PERSPECTIVES AND DISCUSSION

Terminus-associated non-coding RNAs were identified years ago, however, their definition is somewhat confused for researchers. Regarding the nomenclature, the abbreviation "UaRNAs" has been used to indicate "upstream antisense RNAs" and sometimes "3' UTR-associated RNAs." In terms of timing, upstream antisense RNAs (UaRNAs) were reported before 3' UTR-associated RNAs (uaRNAs) (Chiu et al., 2006). uaRNAs were then defined according to their specific genomic location within mRNA 3' UTRs (Mercer et al., 2011). However, UaRNAs were also later discovered and studied (Flynn et al., 2011). Therefore, the abbreviation "uaRNAs" has been given different meanings in separate studies, possibly causing confusion (Chiu et al., 2006; Flynn et al., 2011; Mercer et al., 2011; Lu et al., 2015; Ogami et al., 2017; Yu et al., 2018). To some degree, TTSa-RNAs and TANRs broadly belong to the same class of TASRs. Indeed, when TASRs were first and systemically described, no identifiable patterns, such as genomic locations, lengths, and subcellular localizations, were unraveled (Kapranov et al., 2007). In contrast, TTSa-RNAs enriched in small RNA libraries of AGO1/2 immunoprecipitates are located before the cleavage sites of mRNAs with restricted lengths (approximately 23 nt) and exhibit nuclear localization (Valen et al., 2011;

Laudadio et al., 2018). Moreover, unlike other TANRs, TANRs are a well-known group of snoRNAs (Leng et al., 2014). Regarding CUTs and SUTs, although no clear partition between CUTs and SUTs exists, some ncRNAs defined as CUTs have been redefined as SUTs (Jacquier, 2009; Neil et al., 2009; Xu et al., 2009). Recently, a uniform annotation system for transcript boundaries has been proposed. This annotation is based on their genomic positions and sequence lengths, and provides suggestions for additional classifications of TANRs, for instance according to their biogenesis pathways, modes of action, and biological outputs (Yu et al., 2018). However, as more and diverse TANRs are found in other eukaryotes, a new, more elaborate nomenclature for TANR classification should be proposed, including detailed information on their genomic location, originating strand, biogenesis pathway, and functions.

The discovery of novel transcripts around annotated transcripts also challenges the concept of gene (Gerstein et al., 2007; Gingeras, 2007). Indeed, not only mRNAs but also lncRNAs can generate functional TANRs. A well-known example is *MALAT1*-associated small cytoplasmic RNA (mascRNA), generated from the nascent lncRNA metastasis associated lung adenocarcinoma transcript 1 (*MALAT1*). MascRNA is located at the 3' end of mature *MALAT1*, and its maturation is dependent on RNase P (Wilusz et al., 2008). Functional studies of mascRNA found that this ncRNA is involved in cardiovascular innate immunity (Gast et al., 2016). Surprisingly, mascRNA could function as a translational enhancer when placed downstream of *cGFP in vivo* (Wilusz et al., 2012). Studies of the function and biogenesis of mascRNA suggested that TANRs originating from lncRNAs also play an important role in regulating gene expression. Furthermore, several studies have found that ends of both some mRNAs and certain lncRNAs contained conserved secondary structures that might generate TANRs (Kertesz et al., 2010; Nguyen et al., 2016; Yu et al., 2017). Hence, the possible presence of TANRs should not be ignored in either protein-coding or non-protein-coding loci.

Although TANRs can derive from different pathways, their biogenesis might involve the cross-talk of several regulatory mechanisms. For instance, transcription and posttranscriptional processing are important steps of the maturation of TANRs. Furthermore, the carboxy-terminal domain (CTD) of RNAPII is important for coupling mRNA transcription and processing (McCracken et al., 1997; Proudfoot et al., 2002; Ahn et al., 2004; Bentley, 2005). Indeed, by interacting with splicing and 3' cleavage factors, RNAPII couples transcription, splicing, and cleavage of mRNA precursors (Ahn et al., 2004). Meanwhile, terminal sites are associated with pausing of RNA polymerase (Schwalb et al., 2016). Thus, whether TANRs mature during a coupled process of transcription and posttranscriptional cleavage or they are derived from RNAPII backtracking remains unknown. Thus, new methods for detecting nascent RNAs or the use of mutants in mRNA 3' end maturation pathways may shed some light on TANR biogenesis (Wissink et al., 2019; Furlan et al., 2020).

Given the heterogeneity of TANRs, unraveling their functions has become one of the most basic and pressing issues. 3' UTRs usually harbor critical elements for gene expression, such

as miRNA response elements (MREs). Therefore, TANRs that contain MREs may act as miRNA sponges, thus protecting the corresponding mRNAs from translational repression or degradation. For instance, uaRNA *FLJ11812* functions as a ceRNA by sponging *miR-4459* from its target mRNAs, thereby providing a novel direction for functional studies (Ge et al., 2014; Lu et al., 2015). Furthermore, the formation of gene loops juxtaposing the promoter and terminator has been reported in several organisms, and gene looping is thought to mediate long-distance transcriptional regulation (Bratkowski et al., 2018). However, it is unclear whether TANR-mediated gene looping is required for guiding DNA methylation, mRNA processing, or other processes. Nevertheless, the occurrence of miRNA sponging and gene looping provides novel directions for functional studies of sense TANRs. As for antisense ncRNAs, the discovery of chromatin-based transcription interference also suggested a new mechanism of TANR function (Gill et al., 2020).

Once a TANR is discovered, it is challenging to know how to study its function. Basic information, such as the abundance of related mRNAs, the secondary structure, and the subcellular localization of TANRs, aids in understanding their possible functions. For detailed functional studies, induced upregulation and downregulation of TANRs represent an appropriate strategy for primary functional studies. To achieve upregulation, overexpression or synthesis of certain TANRs represents available methods. However, for most TANRs overlapping with certain 3' UTRs that harbor important regulatory elements, some technical issues need to be overcome to eliminate the potential impact of induced downregulation on the upstream transcripts. Currently, siRNA screens and the application of CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 to delete certain DNA regions provide useful tools for functional annotation of TANRs in a native context (Zhao et al., 2017).

In summary, the discovery of TANRs in different eukaryotes suggested that they are abundant and conserved. Moreover, studies of the biogenesis and functions of TANRs indicated that they can play important roles in different cellular activities. However, since TANRs represent a novel group of ncRNAs, their biogenesis and functions still require further research. As more information about different TANRs is being reported, their involvement in the regulation of gene expression is due to be unfolded in full and presents one more intriguing observation of the versatility of RNA function.

AUTHOR CONTRIBUTIONS

FX and X-ML contributed to the conception of the study. W-JN and X-ML wrote the manuscript. W-JN, FX, and X-ML discussed and improved the revised manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China (81760506).

REFERENCES

- Ahn, S. H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol. Cell* 13, 67–76. doi: 10.1016/s1097-2765(03)00492-1
- Amack, J. D., Paguio, A. P., and Mahadevan, M. S. (1999). Cis and trans effects of the myotonic dystrophy (DM) mutation in a cell culture model. *Hum. Mol. Genet.* 8, 1975–1984. doi: 10.1093/hmg/8.11.1975
- Andreassi, C., and Riccio, A. (2009). To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol.* 19, 465–474. doi: 10.1016/j.tcb.2009.06.001
- Ayupé, A. C., Tahira, A. C., Camargo, L., Beckedorff, F. C., Verjovski-Almeida, S., and Reis, E. M. (2015). Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. *RNA Biol.* 12, 877–892. doi: 10.1080/15476286.2015.1062960
- Bachellerie, J. P., Cavaille, J., and Huttenhofer, A. (2002). The expanding snoRNA world. *Biochimie* 84, 775–790. doi: 10.1016/s0300-9084(02)01402-5
- Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363.
- Bentley, D. L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* 17, 251–256. doi: 10.1016/j.ccb.2005.04.006
- Bratkowski, M., Unarta, I. C., Zhu, L., Shubbar, M., Huang, X., and Liu, X. (2018). Structural dissection of an interaction between transcription initiation and termination factors implicated in promoter-terminator cross-talk. *J. Biol. Chem.* 293, 1651–1665. doi: 10.1074/jbc.m117.811521
- 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
- Carbonell, A. (2017). Plant ARGONAUTES: features, functions, and unknowns. *Methods Mol. Biol.* 1640, 1–21. doi: 10.1007/978-1-4939-7165-7_1
- Chiu, A., Wu, X., Subtelny, A., and Sharp, P. (2006). Characterizing polyadenylated uRNAs suggests a potential role for Pabpn1. *Environ. Prot. Chem. Ind.* 29:S1. doi: 10.1176/appi.pn.2014.3a2
- Clark, M. B., Amaral, P. P., Schlesinger, F. J., Dinger, M. E., Taft, R. J., Rinn, J. L., et al. (2011). The reality of pervasive transcription. *PLoS Biol.* 9:e1000625. doi: 10.1371/journal.pbio.1000625
- Daugaard, I., and Hansen, T. B. (2017). Biogenesis and function of ago-associated RNAs. *Trends Genet.* 33, 208–219. doi: 10.1016/j.tig.2017.01.003
- Denti, M. A., Viero, G., Provenzani, A., Quattrone, A., and Macchi, P. (2013). mRNA fate: life and death of the mRNA in the cytoplasm. *RNA Biol.* 10, 360–366. doi: 10.4161/rna.23770
- Ding, N., Qu, H., and Fang, X. (2014). [The ENCODE project and functional genomics studies]. *Yi Chuan* 36, 237–247.
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108.
- Duempelmann, L., Skribbe, M., and Buhler, M. (2020). Small RNAs in the transgenerational inheritance of epigenetic information. *Trends Genet.* 36, 203–214. doi: 10.1016/j.tig.2019.12.001
- Flynn, R. A., Almada, A. E., Zamudio, J. R., and Sharp, P. A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10460–10465. doi: 10.1073/pnas.1106630108
- Forrest, A. R., and Carninci, P. (2009). Whole genome transcriptome analysis. *RNA Biol.* 6, 107–112. doi: 10.4161/rna.6.2.7931
- Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., et al. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* 39, 1033–1037. doi: 10.1038/ng2079
- Furlan, M., Tanaka, I., Leonardi, T., De Pretis, S., and Pelizzola, M. (2020). Direct RNA sequencing for the study of synthesis, processing, and degradation of modified transcripts. *Front. Genet.* 11:394. doi: 10.3389/fgene.2020.00394
- Gast, M., Schroen, B., Voigt, A., Haas, J., Kuehl, U., Lassner, D., et al. (2016). Long noncoding RNA MALAT1-derived mascRNA is involved in cardiovascular innate immunity. *J. Mol. Cell Biol.* 8, 178–181. doi: 10.1093/jmcb/mjw003
- Ge, D., Han, L., Huang, S., Peng, N., Wang, P., Jiang, Z., et al. (2014). Identification of a novel MTOR activator and discovery of a competing endogenous RNA regulating autophagy in vascular endothelial cells. *Autophagy* 10, 957–971. doi: 10.4161/auto.28363
- Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., et al. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Res.* 17, 669–681. doi: 10.1101/gr.6339607
- Gill, J. K., Maffioletti, A., Garcia-Moliner, V., Stutz, F., and Soudet, J. (2020). Fine chromatin-driven mechanism of transcription interference by antisense noncoding transcription. *Cell Rep.* 31:107612. doi: 10.1016/j.celrep.2020.107612
- Gingeras, T. R. (2007). Origin of phenotypes: genes and transcripts. *Genome Res.* 17, 682–690. doi: 10.1101/gr.6525007
- Jacquier, A. (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat. Rev. Genet.* 10, 833–844. doi: 10.1038/nrg2683
- Jenny, A., Hachet, O., Závorszky, P., Cyrklaff, A., and Weston, M. D. (2006). A translation-independent role of oskar RNA in early *Drosophila* oogenesis. *Development* 133, 2827–2833. doi: 10.1242/dev.02456
- Jensen, T. H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. *Mol. Cell* 52, 473–484. doi: 10.1016/j.molcel.2013.10.032
- Jia, J., Yao, P., Arif, A., and Fox, P. L. (2013). Regulation and dysregulation of 3'UTR-mediated translational control. *Curr. Opin. Genet. Dev.* 23, 29–34. doi: 10.1016/j.gde.2012.12.004
- Jiang, H., Kim, D., and Morris, K. V. (2007). Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12422–12427. doi: 10.1073/pnas.0701635104
- Joshua-Tor, L., and Hannon, G. J. (2011). Ancestral roles of small RNAs: an Ago-centric perspective. *Cold Spring Harb. Perspect. Biol.* 3:a003772. doi: 10.1101/cshperspect.a003772
- 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., Ozsolak, 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
- Kertesz, M., Wan, Y., Mazor, E., Rinn, J. L., Nutter, R. C., Chang, H. Y., et al. (2010). Genome-wide measurement of RNA secondary structure in yeast. *Nature* 467, 103–107. doi: 10.1038/nature09322
- Kim, V. N., Han, J., and Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139. doi: 10.1038/nrm2632
- Laudadio, I., Formichetti, S., Gioiosa, S., Klironomos, F., Rajewsky, N., Macino, G., et al. (2018). Characterization of transcription termination-associated RNAs: new insights into their biogenesis, tailing, and expression in primary tumors. *Int. J. Genomics* 2018:1243858.
- Leng, X. M., Diao, L. T., Li, B., Bi, Y. Z., 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
- Li, Y., Liu, X., Huang, L., Guo, H., and Wang, X. J. (2010). Potential coexistence of both bacterial and eukaryotic small RNA biogenesis and functional related protein homologs in Archaea. *J. Genet. Genomics* 37, 493–503. doi: 10.1016/s1673-8527(09)60069-2
- Liu, J., Wang, H., and Chua, N. H. (2015). Long noncoding RNA transcriptome of plants. *Plant Biotechnol. J.* 13, 319–328. doi: 10.1111/pbi.12336
- Lu, W., Han, L., Su, L., Zhao, J., Zhang, Y., Zhang, S., et al. (2015). A 3'UTR-associated RNA, FLJ11812 maintains stemness of human embryonic stem cells by targeting miR-4459. *Stem Cells Dev.* 24, 1133–1140. doi: 10.1089/scd.2014.0353
- Lu, Z., and Lin, Z. (2019). Pervasive and dynamic transcription initiation in *Saccharomyces cerevisiae*. *Genome Res.* 29, 1198–1210. doi: 10.1101/gr.245456.118
- Ma, X., Han, N., Shao, C., and Meng, Y. (2017). Transcriptome-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
- Maida, Y., Yasukawa, M., and Masutomi, K. (2016). De novo RNA synthesis by RNA-dependent RNA polymerase activity of telomerase reverse transcriptase. *Mol. Cell Biol.* 36, 1248–1259. doi: 10.1128/mcb.01021-15

- Mattick, J. S. (2003). Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays* 25, 930–939. doi: 10.1002/bies.10332
- Mayr, C. (2017). Regulation by 3'-untranslated regions. *Annu. Rev. Genet.* 51, 171–194. doi: 10.1146/annurev-genet-120116-024704
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., et al. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385, 357–361. doi: 10.1038/385357a0
- Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.* 14, 447–459. doi: 10.1038/nrg3462
- Mercer, T. R., Wilhelm, D., Dinger, M. E., Solda, G., Korbje, D. J., Glazov, E. A., et al. (2011). Expression of distinct RNAs from 3' untranslated regions. *Nucleic Acids Res.* 39, 2393–2403. doi: 10.1093/nar/gkq1158
- Moraes, F., and Goes, A. (2016). A decade of human genome project conclusion: scientific diffusion about our genome knowledge. *Biochem. Mol. Biol. Educ.* 44, 215–223. doi: 10.1002/bmb.20952
- Neil, H., Malabat, C., D'aubenton-Carafa, Y., Xu, Z., Steinmetz, L. M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 457, 1038–1042. doi: 10.1038/nature07747
- Nevers, A., Doyen, A., Malabat, C., Neron, B., Kergrohen, T., Jacquier, A., et al. (2018). Antisense transcriptional interference mediates condition-specific gene repression in budding yeast. *Nucleic Acids Res.* 46, 6009–6025. doi: 10.1093/nar/gky342
- Nguyen, T. C., Cao, X., Yu, P., Xiao, S., Lu, J., Biase, F. H., et al. (2016). Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. *Nat. Commun.* 7:12023.
- Ni, W.-J., and Leng, X.-M. (2015). Dynamic miRNA-mRNA paradigms: new faces of miRNAs. *Biochem. Biophys. Rep.* 4, 337–341. doi: 10.1016/j.bbrep.2015.10.011
- Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J. J., et al. (2017). An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* 31, 1257–1271. doi: 10.1101/gad.302604.117
- Pánek, J., Kolář, M., Herrmannová, A., and Valášek, L. S. (2016). A systematic computational analysis of the rRNA-3' UTR sequence complementarity suggests a regulatory mechanism influencing post-termination events in metazoan translation. *RNA* 22, 957–967. doi: 10.1261/rna.056119.116
- Preker, P., Almvig, K., Christensen, M. S., Valen, E., Mapendano, C. K., Sandelin, A., et al. (2011). PROMoter uPstream transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. *Nucleic Acids Res.* 39, 7179–7193. doi: 10.1093/nar/gkr370
- Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mRNA processing with transcription. *Cell* 108, 501–512. doi: 10.1016/s0092-8674(02)00617-7
- Rastinejad, F., and Blau, H. M. (1993). Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation. *Cell* 72, 903–917. doi: 10.1016/0092-8674(93)90579-f
- Saeed, S., Zahra, V. A., Sharif, M., Ali, A., Javad, D. S., and Sadegh, A. J. (2020). Small regulatory noncoding RNAs in *Drosophila melanogaster*: biogenesis and biological functions. *Brief. Funct. Genomics* 19, 309–323. doi: 10.1093/bfpg/ela005
- Schwab, B., Michel, M., Zacher, B., Fruhauf, K., Demel, C., Tresch, A., et al. (2016). TT-seq maps the human transient transcriptome. *Science* 352, 1225–1228. doi: 10.1126/science.aad9841
- Scott, M. S., and Ono, M. (2011). From snoRNA to miRNA: dual function regulatory non-coding RNAs. *Biochimie* 93, 1987–1992. doi: 10.1016/j.biochi.2011.05.026
- 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
- Taft, 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.
- Thomas, Q. A., Ard, R., Liu, J., Li, B., Wang, J., Pelechano, V., et al. (2020). Transcript isoform sequencing reveals widespread promoter-proximal transcriptional termination in *Arabidopsis*. *Nat. Commun.* 11:2589.
- 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. *BMC Genomics* 18:285. doi: 10.1186/s12864-017-3662-1
- 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. Struct. Mol. Biol.* 18, 1075–1082. doi: 10.1038/nsmb.2091
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci.* 13, 350–358. doi: 10.1016/j.tplants.2008.04.007
- 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
- Wickens, M., Anderson, P., and Jackson, R. J. (1997). Life and death in the cytoplasm: messages from the 3' end. *Curr. Opin. Genet. Dev.* 7, 220–232. doi: 10.1016/s0959-437x(97)80132-3
- Willingham, A. T., Dike, S., Cheng, J., Manak, J. R., Bell, I., Cheung, E., et al. (2006). Transcriptional landscape of the human and fly genomes: nonlinear and multifunctional modular model of transcriptomes. *Cold Spring Harb. Symp. Quant. Biol.* 71, 101–110. doi: 10.1101/sqb.2006.71.068
- Wilusz, J. E., Freier, S. M., and Spector, D. L. (2008). 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* 135, 919–932. doi: 10.1016/j.cell.2008.10.012
- Wilusz, J. E., Injapbaptiste, C. K., Lu, L. Y., Kuhn, C. D., Joshua-Tor, L., and Sharp, P. A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* 26, 2392–2407. doi: 10.1101/gad.204438.112
- Wissink, E. M., Vihervaara, A., Tippens, N. D., and Lis, J. T. (2019). Nascent RNA analyses: tracking transcription and its regulation. *Nat. Rev. Genet.* 20, 705–723. doi: 10.1038/s41576-019-0159-6
- 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
- Yamamoto, N., Agata, K., Nakashima, K., and Imamura, T. (2016). Bidirectional promoters link cAMP signaling with irreversible differentiation through promoter-associated non-coding RNA (pancRNA) expression in PC12 cells. *Nucleic Acids Res.* 44, 5105–5122. doi: 10.1093/nar/gkw113
- 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.
- Yu, D., Ma, X., Zuo, Z., Wang, H., and Meng, Y. (2018). Classification of transcription boundary-associated RNAs (TBARs) in animals and plants. *Front. Genet.* 9:168. doi: 10.3389/fgene.2018.00168
- 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
- Zhao, W., Siegel, D., Biton, A., Tonqueze, O. L., Zaitun, N., Ahituv, N., et al. (2017). CRISPR-Cas9-mediated functional dissection of 3'-UTRs. *Nucleic Acids Res.* 45, 10800–10810. doi: 10.1093/nar/gkx675

Conflict of Interest: 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 © 2020 Ni, Xie and Leng. 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(s) 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.