The International Journal of Biochemistry & Cell Biology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Review

Life without A tail: New formats of long noncoding RNAs☆

Yang Zhang[a], Li Yang[b], Ling-Ling Chen[a,*]

[a] State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai, China
[b] Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

ARTICLE INFO

Article history:
Received 11 August 2013
Received in revised form
30 September 2013
Accepted 7 October 2013
Available online xxx

Keywords:
nonpolyadenylated RNAs
MALAT1
NEAT1
sno-lncRNAs
circular RNAs
ciRNAs

ABSTRACT

While most long noncoding RNAs (lncRNAs) appear indistinguishable from mRNAs, having 5′ cap structures and 3′ poly(A) tails, recent work has revealed new formats. Rather than taking advantage of the canonical cleavage and polyadenylation for their 3′ end maturation, such lncRNAs are processed and stablized by a number of other mechanisms, including the RNase P cleavage to generate a mature 3′ end, or capped by snoRNP complexes at both ends, or by forming circular structures. Importantly, such lncRNAs have also been implicated in gene expression regulation in mammalian cells. Here, we highlight recent progress in our understanding of the biogenesis and function of lncRNAs without a poly(A) tail.

This paper is part of a directed issue entitled: The Non-coding RNA Revolution.

Contents

1. Introduction

In eukaryotic cells, nascent pre-mRNA transcripts undergo multiple co-/post-transcriptional processing and modification events during their maturation, which are tightly coupled and precisely regulated. The 3′ end maturation of a nascent transcript is critical for allowing the release of the RNA from the transcription template and for assuring the functionality of a mature mRNA. With the exception of replication-dependent histone mRNAs, the mechanism of 3′ end processing of nearly all RNA polymerase II (Pol II) transcribed mRNAs in eukaryotic cells involves cleavage and polyadenylation of nascent transcripts. The cleavage/polyadenylation machinery first recognizes the AAUAAA hexanucleotide (or some variants), often together with a downstream G/U rich sequence present in nascent transcripts, resulting in endonucleolytic cleavage of the pre-mRNA by CPSF-73 (Mandel et al., 2006). A poly(A) tail of up to 200 to 250 adenosines is subsequently added by poly(A) polymerase to the 3′-end of the transcript in a non-templated fashion (reviewed in Shatkin and Manley, 2000; Richard and Manley, 2009).

Recent studies revealed that alternative cleavage and polyadenylation (APA) occur at more than 70% of all human

☆ This paper is part of a directed issue entitled: The Non-coding RNA Revolution.

* Corresponding author. Tel.: +86 2154921021.
E-mail address: linglingchen@sibcb.ac.cn (L.-L. Chen).

1357-2725/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biocel.2013.10.009

genes (Tian et al., 2005; Hoque et al., 2013), leading to isoforms of a mRNA with variable coding potentials and/or 3′ untranslated regions (UTRs), and subsequently, different stability and subcellular localization (reviewed in Tian and Manley, 2013). As such, the selection of each 3′ end cleavage site is under tight control and the aberrant usage has a great impact on cellular regulation (reviewed in Tian and Manley, 2013). Thus, it is becoming increasingly evident that the proper processing of 3′ end is crucial to the functionality of a mature mRNA. Interestingly, although it is generally believed that replication-dependent histone mRNAs are the only cellular mRNAs without a poly(A) tail, several large scale transcriptomic studies have found that mature mRNAs with short poly(A) tails of less than 30 nucleotides also exist in mammalian cells (Gu et al., 1999; Cheng et al., 2005; Wu et al., 2008; Yang et al., 2011a), albeit the mechanism and function of the appearance of the short poly(A) tail remain unclear.

As the great majority of eukaryotic mRNAs end in a poly(A) tail, transcriptome analyses using tiling arrays and deep sequencing (mRNA-seq) in earlier studies typically involved the enrichment of poly(A)+ RNAs by oligo(dT) selection (Pan et al., 2008; Wilhelm et al., 2008). In addition to the fine mapping of mRNAs, this type of analyses surprisingly revealed the widespread transcription of long noncoding RNAs (lncRNAs) in mammalian cells. These lncRNAs include natural antisense transcripts (NATs) (Katayama et al., 2005) and thousands of long noncoding RNAs transcribed from intergenic regions (lincRNAs) (Okazaki et al., 2002; Guttman et al., 2009, 2011; Cabili et al., 2011; Derrien et al., 2012) (Fig. 1). Characteristics that define a lincRNA in a variety of organisms are based on knowlege obtained from features of a mRNA, including chromatin maps and data from methods used to identify Pol II transcription start and elongation regions, such as histone 3 Lys 4 trimethylation in the promoter region and histone 3 Lys 36 trimethylation across the actively transcribed region, and the presence of a poly(A) site for transcription termination (Khalil et al., 2009; Cabili et al., 2011; Ulitsky et al., 2011; Guttman and Rinn, 2012). Thus, at the molecular level, most lincRNAs appear indistinguishable from mRNAs, with 5′ cap structures, 3′ poly(A) tails and exon–exon splice junctions. As it has long been thought that the proper 3′ end processing of a nascent transcript is critical for the fate of a mature mRNA, a poly(A) tail is presumably essential for the stability and functionality of a lncRNA. Although the functions of many lncRNAs remain mysterious, new lines of evidence on several dozens of well-characterized ones have strongly implicated them in a variety of important biological processes (Reviewed in Ulitsky and Bartel, 2013; Rinn and Chang, 2012).

In contrast to the widespread expression of polyadenylated lncRNAs, recent studies revealed that a number of Pol II transcribed lncRNAs, including a few lincRNAs and others originate from different genomic locations, such as enhancer and intron regions, and are processed in alternative ways (Wilusz et al., 2008; Sunwoo et al., 2009; Kim et al., 2010; De Santa et al., 2010; Burd et al., 2010; Yap et al., 2010; Hansen et al., 2013; Salzman et al., 2012; Yin et al., 2012; Hansen et al., 2013; Jeck et al., 2013; Memczak et al., 2013) (Fig. 1). Enhancer RNAs (eRNAs) are less than 2 kb in length, bidirectionally transcribed from enhancer regions, largely lack poly(A) tails, and have very low copy numbers (Kim et al., 2010; De Santa et al., 2010). Although the mechanism of their processing is not yet clear, recent studies revealed that eRNAs are not by-products of transcription but carry out an important role for enhancer-like function (Orom et al., 2010; Kim et al., 2010; Li et al., 2013). For other non-polyadenylated lncRNAs, rather than taking advantage of the canonical cleavage and polyadenylation for their 3′ end maturation, stabilization is achieved by a number of mechanisms, including the RNase P cleavage to generate a mature 3′ end (Brown et al., 2012; Wilusz et al., 2012), or capped by snoRNP complexes at both ends (Yin et al., 2012), or by forming circular structures to protect them from degradation during their life cycle (Hansen et al., 2013; Jeck et al., 2013; Memczak et al., 2013; Zhang et al., 2013). In this review, we will highlight the recent progress of the biogenesis and function of these non-polyadenylated lncRNAs.

2. RNase P cleavage leads to some mature lncRNAs with triple-helical RNA structures at their 3′ ends

RNase P is an endoribonuclease and is best known for its function in tRNA maturation. In bacteria, RNase P has two components: M1 RNA and the single protein C5. While M1 is the catalytic subunit of the enzyme, C5 is involved in enhancing M1's affinity for substrates and for cleavage efficiency (Guerrier-Takada et al., 1983;

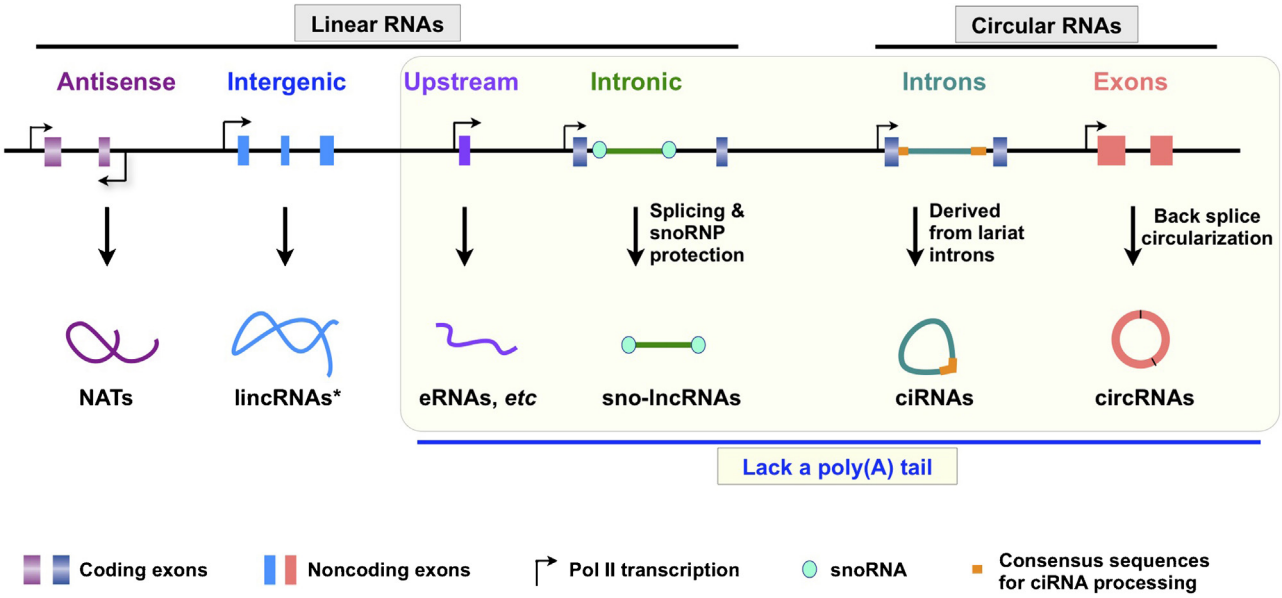

Fig. 1. *Pervasive transcription of various classes of lncRNAs in mammalian genomes.* NATs, natural antisense transcripts; lincRNAs, large intergenic noncoding RNAs; eRNAs, enhancer RNAs; sno-lncRNAs, snoRNA-related lncRNAs; ciRNAs, circular intronic RNAs; circRNAs, circular RNAs. *There are a few reported lincRNAs without a poly(A) tail. See text for details.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

Sun et al., 2006, 2010). In human, the catalytic RNA-Protein (RNP) complex has an H1 RNA moiety and at least ten protein subunits that are associated with H1 (Reiner et al., 2011). H1 RNA alone is catalytically inactive in *in vitro* assays, but exerts its endonucleolytic activity with the support of the Rpp21/Rpp29 heterodimer (Reiner et al., 2011). In addition to its function in tRNA processing and the transcription of small ncRNAs by RNA Polymerases (Marvin et al., 2011), recent studies have revealed that RNase P complex is involved in the 3′ end maturation of some lncRNAs in yeast (Yang and Altman, 2007) as well as of *MALAT1* and *NEAT1* in mammals (Wilusz et al., 2008).

2.1. MALAT1

MALAT1 (metastasis-associated lung adenocarcinoma transcript 1), also known as *NEAT2* (nuclear-enriched abundant transcript 2) (Hutchinson et al., 2007) was first identified as a prognostic marker of survival in early-stage non-small cell lung cancers (Ji et al., 2003). Besides its over-expression in carcinomas, *MALAT1* also exhibits a wide spectrum of expression in normal tissues with relatively higher expression in brain (Ji et al., 2003; Bernard et al., 2010). In contrast to the fact that many known regulatory lncRNAs are less conserved than protein coding genes (Cabili et al., 2011; Derrien et al., 2012), *MALAT1* exhibits a high conservation in mammals (Hutchinson et al., 2007) and other lower vertebrates, such as zebrafish (Ulitsky et al., 2011), suggesting strongly that it has important roles in biological processes.

MALAT1 is one of the most abundant nuclear enriched lncRNAs and is localized to nuclear speckles (Hutchinson et al., 2007; Tripathi et al., 2010; Spector and Lamond, 2011). Careful analyses by Wilusz et al. have revealed that mouse *MALAT1* has two isoforms. The ~7 kb (~8 kb in human) long transcript has conserved polyadenylation signals at the 3′ end and is retained in nucleus (Wilusz et al., 2008). However, the long isoform is present at a very low level that cannot be easily detected (Wilusz et al., 2008), which explains *MALAT1* only showed one band in previous studies (Ji et al., 2003; Lin et al., 2007; Hutchinson et al., 2007). The abundant shorter transcript of *MALAT1* is ~6.7 kb (~7.5 kb in human), which is generated by RNase P processing to cleave off a highly conserved tRNA-like structure from the 3′ end of the nascent *MALAT1* transcript. Cleavage by RNase P generates a stable nuclear-retained *MALAT1* with a short A-rich tract and a tRNA like, 61 nt cytoplasmic mascRNA (*MALAT1*-associated small cytoplasmic RNA) with unknown function (Wilusz et al., 2008) (Fig. 2).

Interestingly, the short A-rich motif at the 3′ end of *MALAT1* is derived from a genomically encoded A-rich tract rather than by being added post-transcriptionally, but it offers an even more stable status for *MALAT1* than many RNAs with canonical poly(A) tails. How is the stability at the end of *MALAT1* is achieved? It turns out that the 3′ end of *MALAT1* contains highly conserved U- and A-rich sequences, which was proposed to form a stable U-A·U triple helical structure (Wilusz et al., 2012; Brown et al., 2012). This triple helical structure constitutes the U-rich motif 2 and the A-rich motif that form Watson–Crick base pair helix while the U-rich motif 1 connects to A-rich motif by Hoogsteen hydrogen bond (Wilusz et al., 2012; Brown et al., 2012) (Fig. 2). A similar triple helical structure at the 3′ end has also been seen in another mammalian lncRNA, *NEAT1* (see below) (Fig. 3), *PAN* lncRNA from Kaposi's sarcoma-associated herpesvirus and in the RNAs from other viruses (Mitton-Fry et al., 2010; Tycowski et al., 2012). More strikingly, when the triple helical structure from *MALAT1* was placed on a GFP reporter, the resulting mRNA could be exported to the cytoplasm and translated, suggesting that this unexpected structure at the 3′ end can support export and translation similar to a poly(A) tail (Wilusz et al., 2012), although it is elusive whether such a U-A·U triple helical structure exists in any endogenous mRNAs.

Although *MALAT1* is exclusively localized to nuclear speckles (Hutchinson et al., 2007; Tripathi et al., 2010), it is not a structural

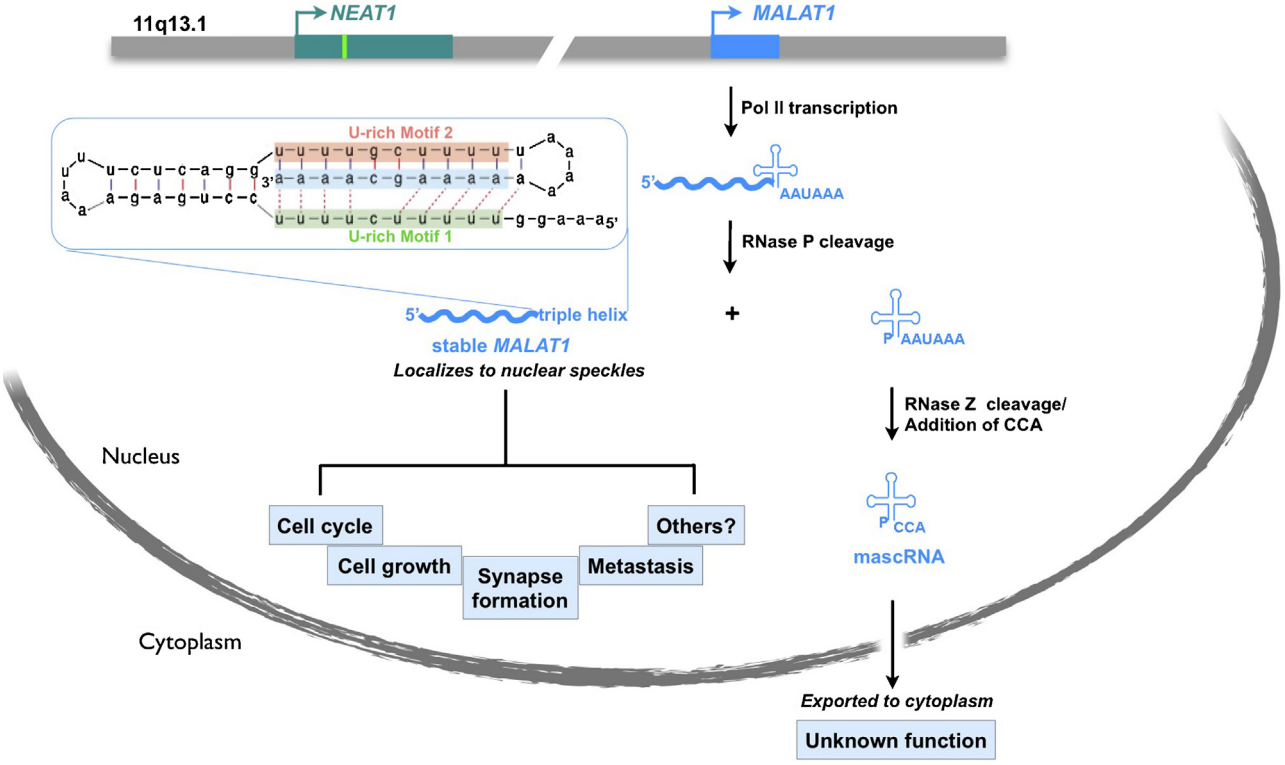

Fig. 2. *RNase P processing of the 3′ end of MALAT1*. The nascent *MALAT1* transcript forms a tRNA-like structure at its 3′ end, which can be recognized and cleaved by RNase P to generate a stable *MALAT1* with a U-A·U triple helical structure at the 3′ end. The U-A·U triple helical structure has been modified from (Wilusz et al., 2012). *MALAT1* plays an important role in several biological processes by regulating alternative splicing of pre-mRNAs and modulating expression of growth control genes.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

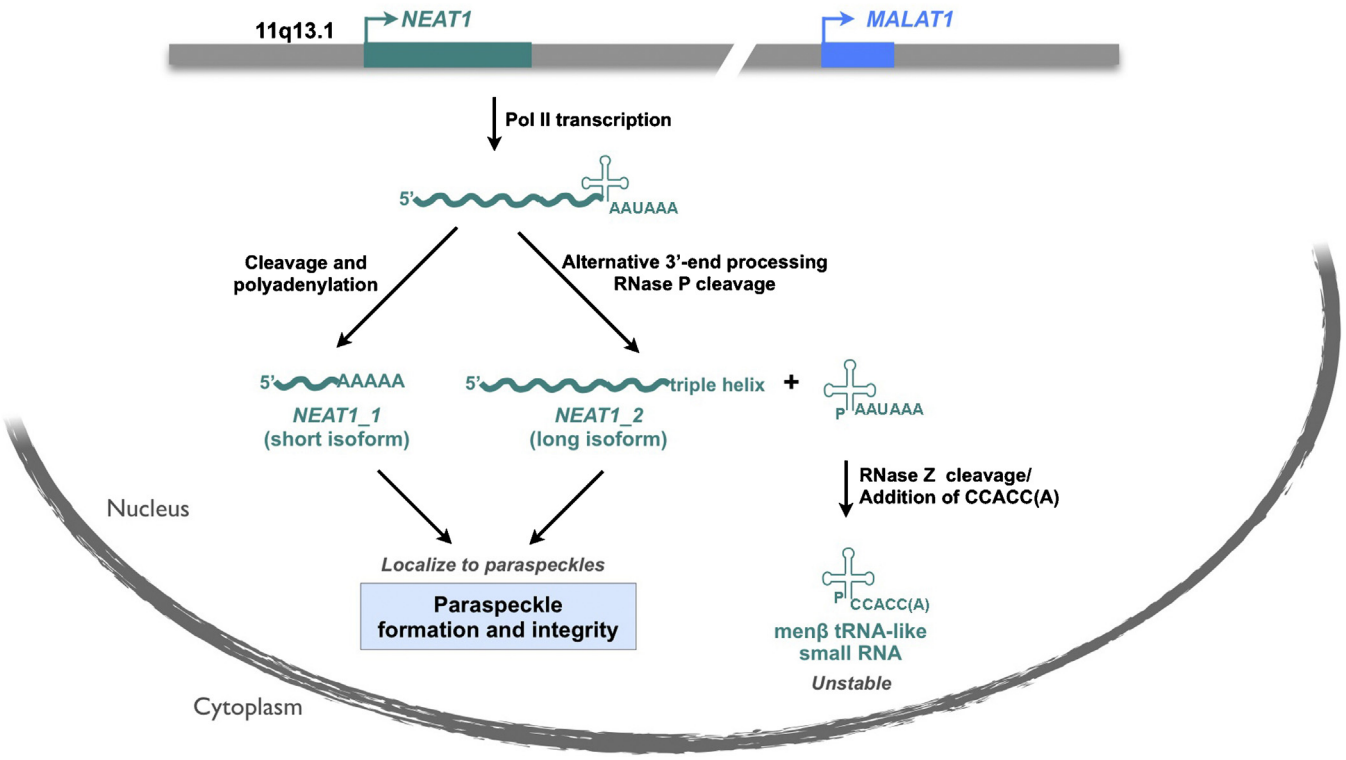

Fig. 3. *Processing of two isoforms of NEAT1. NEAT1* has two isoforms. The short *NEAT1_1* is polyadenylated, while the long *NEAT1_2* is processed by RNase P at its 3′ end (Sunwoo et al., 2009; Wilusz et al., 2011). Although both isoforms are essential for the integrity of paraspeckles, the long isoform is more important for *de novo* paraspeckle assembly.

RNA. Depletion of *MALAT1* did not destroy the integrity of nuclear speckles in both human and mouse cells (Tripathi et al., 2010; Nakagawa et al., 2012; Zhang et al., 2012). In addition, *MALAT1* does not overlap with SC35 domains immediately after mitosis suggesting that it is not necessary for nuclear speckle assembly and that it is more likely to participate in the function of nuclear speckles, such as assembly, modification, and storage of components of the pre-mRNA processing machinery (Hutchinson et al., 2007; Tripathi et al., 2010; Nakagawa et al., 2012; Zhang et al., 2012). At the molecular level, it has been shown that *MALAT1* regulates alternative splicing by modulating the levels of active (phosphorylated) serine/arginine (SR) proteins in HeLa cells (Tripathi et al., 2010). In this scenario, *MALAT1* acts as a decoy that titrates SR splicing factors away from potential endogenous pre-mRNA targets. Depletion of *MALAT1* increases the levels of cellular SR proteins and leads to alternative splicing of a subset of endogenous pre-mRNAs (Tripathi et al., 2010). Furthermore, *MALAT1* was also shown to modulate the recruitment of SR proteins to active transcription sites by a transgene array (Bernard et al., 2010). Moreover, the abundance of *MALAT1* also supports its possible role as a "molecular sponge" for RNA-binding proteins. For example, genome wide studies revealed that *MALAT1* has many binding sites for TDP-43, a protein that binds to pre-mRNAs and regulates alternative splicing (Tollervey et al., 2011; Polymenidou et al., 2011), further pointing out the possibility of *MALAT1* in splicing regulation. However, a discrepancy has been observed from *MALAT1* knockout mice derived cells (Nakagawa et al., 2012; Zhang et al., 2012). Here, depletion of *MALAT1* had no significant effect on splicing factor levels, phosphorylation status, or alternative pre-mRNA splicing (Nakagawa et al., 2012; Zhang et al., 2012). Interestingly, the inactivation of *MALAT1* transcription itself could alter local (*MALAT1* neighboring genes) but not global gene expression, suggesting that the act of *MALAT1* transcription may play a *cis*-regulatory role (Nakagawa et al., 2012; Zhang et al., 2012). Finally, a recent study suggested that *MALAT1* could work with methylation of Polycomb 2 protein (Pc2), a component of Polycomb repressive complex 1 (PRC1) and play a role in the relocation of growth-control genes between nuclear structures and gene activation (Yang et al., 2011b).

Although studies of *MALAT1* loss-of-function genetic models have demonstrated that *MALAT1* is not essential for mouse development under normal laboratory conditions (Eissmann et al., 2012; Zhang et al., 2012; Nakagawa et al., 2012), at the cellular level, recent reports have suggested that *MALAT1* plays a role in synapse formation (Bernard et al., 2010; Polymenidou et al., 2011), cell growth and tumor metastasis (Yang et al., 2011b; Tripathi et al., 2013; Gutschner et al., 2013b). Knockdown of *MALAT1* reduced the synaptic density and altered expression of genes involved in synapse function, suggesting that *MALAT1* likely regulates synapse formation (Bernard et al., 2010). In addition, TDP-43 depletion led to the reduced expression of *MALAT1* and impaired synapse formation (Polymenidou et al., 2011). The strong association of *MALAT1* with TDP-43 (Tollervey et al., 2011) further supported a role for *MALAT1* in synapse regulation.

On the other hand, *MALAT1* has been implicated in the regulation of cell growth and tumor metastasis. First, *MALAT1* likely participates in cell cycle regulation. The level of *MALAT1* is regulated during normal cell cycle progression and knockdown of *MALAT1* led to an increased G1 phase and decreased S phase population (Tripathi et al., 2013). Further study revealed that depletion of *MALAT1* not only led to the activation of p53 and its target genes, but also affected the alternative splicing of an oncogenic transcription factor that is involved in G2/M progression (Tripathi et al., 2013). Second, *MALAT1* is overexpressed in several tumor tissues and its expression is positively correlated with metastasis (reviewed in Gutschner et al., 2013a). Third, *MALAT1* is a potential regulator of metastasis of lung cancer. The *MALAT1* knockout human lung cancer cells were recently developed by using genome editing technology (Gutschner et al., 2011). *MALAT1* depleted cells exhibited a

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

reduction of cell motility in scratch assays and formation of fewer tumor nodules in mouse xenografts (Gutschner et al., 2013b), which were due to an altered expression of metastasis-associated genes resulted from the loss of *MALAT1* (Gutschner et al., 2013b).

2.2. NEAT1

NEAT1 (nuclear enriched abundant transcript 1), also known as *Men (multiple endocrine neoplasia) ε/β* or *VINC* (virus-inducible ncRNA), is another broadly expressed and highly abundant lncRNA in mammalian cell nuclei (Saha et al., 2006; Hutchinson et al., 2007; Clemson et al., 2009; Sunwoo et al., 2009). *NEAT1* is transcribed by Pol II from a region 58 kb upstream of *MALAT1*. Unlike the highly conserved *MALAT1*, *NEAT1* has only two conserved segments between human and mouse (Hutchinson et al., 2007). Although the primary sequence is not well conserved, *NEAT1* and its mouse homologue both localize to paraspeckles and play a role in regulating gene expression by sequestering mRNAs containing inverted repeats within the nucleus (Clemson et al., 2009; Sunwoo et al., 2009; Chen and Carmichael, 2009; Mao et al., 2011a; for reviews, see Fox and Lamond, 2010; Bond and Fox, 2009; Mao et al., 2011b).

There are two isoforms of *NEAT1*. While the abundant short isoform, *NEAT1_1* (3.7 kb in human) or *Menε* (3.2 kb in mouse), is polyadenylated (Hutchinson et al., 2007), the less abundant long isoform, *NEAT1_2* (23 kb in human) or *Menβ* (20 kb in mouse), is non-polyadenylated (Fig. 3). These two isoforms share the same transcriptional start site, but their 3′ ends are processed by distinct mechanisms: canonical polyadenylation (Naganuma et al., 2012) or RNase P cleavage (Sunwoo et al., 2009; Wilusz et al., 2011). *NEAT1_2* is processed by RNase P with a similar mechanism to that of *MALAT1* (Figs. 2 and 3). RNase P cleavage leads to the formation of a mature 3′ end of *NEAT1_2* (Sunwoo et al., 2009; Wilusz et al., 2008, 2011), which is protected by a triple helical structure (Wilusz et al., 2012; Brown et al., 2012). Components of the CFIm complex, such as CPSF6 and NUDT21 have been recently shown to localize to paraspeckles and to be involved in the 3′ end processing of *NEAT1_1*, confirming the involvement of the canonical polyadenylation in *NEAT1_1* maturation (Naganuma et al., 2012). Interestingly, the expression level of *NEAT1_1* and *NEAT1_2* is likely controlled by HNRNPK, which is a newly identified paraspeckle protein and interferes with the CFIm complex cleavage to keep *NEAT1_2* synthesis low (Naganuma et al., 2012).

Paraspeckles, located adjacent to nuclear speckles, were identified in 2002 by a proteomic study aimed at characterizing human nucleoli (Fox et al., 2002). Paraspeckles were originally identified to contain several well-studied Drosophila Behavior and Human Splicing (DBHS) family proteins including PSP1α, $p54^{nrb}$ and PSF (Fox et al., 2002, 2005; Bond and Fox, 2009). Blocking transcription or depleting RNAs both led to a dramatic decrease in the size and number of paraspeckles, indicating that paraspeckles depend on both RNA and on Pol II transcription for their integrity (Fox et al., 2002, 2005; Prasanth et al., 2005). The first paraspeckle RNA was identified in 2005, named *CTN-RNA*, which is transcribed from mouse cationic amino acid transporter 2 (mCAT2) gene. In the nucleus, double stranded RNAs (dsRNAs) serve as efficient substrates for adenosine (A)-to-inosine (I) A-to-I editing by ADAR enzymes. Editing of long dsRNA duplexes can lead to the deamination of up to 50% of the adenosines on each strand of the RNA duplexes. Duplex RNA structures are commonly formed through base-pairing of inverted repeat regions within transcripts. The 3′ UTR of *CTN-RNA* contains such elements for A-to-I editing and was shown to be nuclear-retained and at least partially localized to paraspeckles (Prasanth et al., 2005). Similarly, human mRNAs containing inverted *Alu* elements in their 3′ UTRs are also associated with paraspeckles (Chen et al., 2008; Chen and Carmichael, 2009). Thus, although the function of paraspeckles is not yet completely clear, these studies have suggested that they could be involved in gene expression regulation by trapping edited RNAs within the nucleus.

Knockdown of *CTN-RNA* did not disrupt paraspeckles, indicating that some other RNA serves as an organizer (Prasanth et al., 2005). In 2009, several groups independently identified *NEAT1* lncRNAs as *bona fide* paraspeckle structural RNAs (Sunwoo et al., 2009; Clemson et al., 2009; Sasaki et al., 2009; Chen and Carmichael, 2009). Unlike *CTN-RNA*, *NEAT1* is not A-to-I hyperedited, consistent with its architectural role in paraspeckles (Clemson et al., 2009). Biochemical analyses demonstrated that *NEAT1* and *Men ε/β* transcripts are associated with paraspeckle DBHS proteins (Sunwoo et al., 2009; Sasaki et al., 2009; Clemson et al., 2009; Chen and Carmichael, 2009), indicating that they might serve as platforms for the recruitment of the multifunctional DBHS proteins. In addition, depletion of *NEAT1* (Clemson et al., 2009; Chen and Carmichael, 2009) or *Menε/β* (Sunwoo et al., 2009; Sasaki et al., 2009) led to the loss of paraspeckles and the nuclear localization of PSP1α/$p54^{nrb}$ became more uniformly distributed, while overexpression of *Menε* RNA increased numbers of paraspeckles in both NIH3T3 and HeLa cells (Clemson et al., 2009), suggesting that *NEAT1* and *Men ε/β* are essential paraspeckle components and are functionally conserved.

Detailed studies have suggested that *NEAT1_2* (or *Men β*) is more essential for *de novo* paraspeckle assembly (Sunwoo et al., 2009; Sasaki et al., 2009; Naganuma et al., 2012). For example, transient expression of *NEAT1_2* but not *NEAT1_1* reassembled paraspeckles in *NEAT1* knockout MEFs, suggesting that *NEAT1_2* is the essential component for *de novo* paraspeckle formation (Naganuma et al., 2012). Furthermore, an electron microscopic study revealed that the common *NEAT1* region and *NEAT1_2* 3′-end region were located at the paraspeckle periphery, whereas the central sequences of *NEAT1_2* were enriched in the paraspeckle interior, further indicating *NEAT1_2* is important for the maintenance of paraspeckle integrity (Souquere et al., 2010). Moreover, in some mouse cells, such as parietal cells, due to lack of *Menβ* expression, *Menε* and paraspeckle proteins are distributed throughout the nucleoplasm instead of forming paraspeckle-like punctations (Nakagawa et al., 2011). Finally, to understand the role of *NEAT1* lncRNAs in paraspeckle assembly, a live-cell imaging system was developed that allows for the inducible transcription of *NEAT1* lncRNAs and direct visualization of paraspeckle proteins (Mao et al., 2011a). This study demonstrated that both the active transcription of *NEAT1* and the *NEAT1* lncRNAs regulate paraspeckle maintenance and dynamics (Mao et al., 2011a). While many recently identified lncRNAs localize to specific but uncharacterized nuclear bodies (Sone et al., 2007; Yin et al., 2012), studies on *NEAT1* have suggested one possibility of the importance of such scaffold RNA in the maintenance of nuclear structures.

Although *NEAT1* is broadly expressed in many tissues in adult animals, it is also differentially expressed during mouse C2C12 myoblast (Sunwoo et al., 2009) or human embryonic stem (ES) cells differentiation (Chen and Carmichael, 2009), suggesting a regulatory role of *NEAT1*. Surprisingly, *NEAT1* knockout mice were viable and fertile, with no obvious phenotypic defects (Nakagawa et al., 2011). This raised the question of whether *NEAT1* and paraspeckles are functional. However, since these knockout mice were fed under laboratory growth conditions, obvious phenotypes might be not visualizable until certain stress conditions were present. Current understanding of paraspeckle-related RNAs supports a model where such nuclear structures can be functional during stress. First, *NEAT1* was originally identified by its up-regulation during virus infection in mice (Saha et al., 2006). Second, under cellular stress, the paraspeckle-retained *CTN-RNA* is post-transcriptionally cleaved to rapidly produce protein-coding *mCAT2* mRNA, which is exported efficiently to the cytoplasm for translation, suggesting a role of paraspeckles in regulating gene expression upon physiologic

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

stress (Prasanth et al., 2005). Third, such desequestration might be a general mechanism of paraspeckles under stress, hinted by the different nuclear/cytoplasmic localizations of isoforms of multiple inverted *Alu* elements containing mRNAs in human cells (Chen et al., 2008; Chen and Carmichael, 2009), and knockdown of *NEAT1* in HeLa cells results both in loss of paraspeckles and in enhanced nucleocytoplasmic export of such mRNAs (Chen and Carmichael, 2009). Thus, the systematic determination of paraspeckle-retained RNAs before and after stress treatments will be of great interest to provide new insights into the regulatory role of *NEAT1* and paraspeckles.

3. snoRNA-ended lncRNAs from excised introns

While the majority of known lncRNAs are transcribed from intergenic regions, recent studies have found new formats of lncRNAs derived from excised introns. Intronic sequences account for over 20% of the human genome and provide yet another source to generate new forms of lncRNAs that lack both 5′ cap structures and 3′ poly(A) tails. Although it is generally believed that most introns or intron fragments are unstable (Rodriguez-Trelles et al., 2006), there are exceptions. For instance, there are several hundred known small nucleolar RNAs (snoRNAs) and the great majority of these are encoded within introns (Filipowicz and Pogacic, 2002). SnoRNAs are a family of conserved nuclear RNAs (about 70–200 nt) that are usually concentrated in Cajal bodies or nucleoli where they either function in the modification of small nuclear RNAs (snRNAs) or ribosomal RNAs (rRNAs), or participate in the processing of rRNAs during ribosome subunit maturation (Boisvert et al., 2007; Kiss, 2001; Matera et al., 2007). SnoRNAs are processed from excised and debranched introns by exonucleolytic trimming (Tycowski et al., 1993; Kiss and Filipowicz, 1995) and carry out their functions in complex with specific protein components by forming ribonucleoprotein complexes (snoRNPs) (Kiss, 2001) (Fig. 4A, in green shadow). There are two main classes of snoRNAs: box C/D snoRNAs and box H/ACA snoRNAs, both of which serve as guide RNAs with complementarity to specific target sequences.

The exploration of non-polyadenylated and ribosomal RNAs depleted "poly(A)-/ribo-" transcripts in human transcriptomes identified many excised introns longer than 200 nucelotides that could accumulate to high levels (Yang et al., 2011a), leading to the question that how these excised introns escape from exonucleolytic trimming after splicing. Detailed studies revealed one mechanism for the processing of such excised introns. Introns containing two snoRNAs are processed from their ends by the snoRNP machinery after splicing and the internal intronic sequences between snoRNAs are not removed, leading to the accumulation of lncRNAs with snoRNA ends (*sno-lncRNAs*) (Yin et al., 2012) (Fig. 4A, in blue shadow). Such molecules lack both 5′ cap structures and 3′ poly(A) tails, but are rather capped with snoRNAs at both ends. In addition, *sno-lncRNAs* are widely expressed in cells and tissues and can be produced by either box C/D or box H/ACA snoRNAs in human genome (Yin et al., 2012).

Importantly, the genomic region encoding five abundant *sno-lncRNAs* (15q11–q13) is specifically deleted in Prader–Willi Syndrome (PWS). PWS is a multiple system disorder characterized by global developmental delay, mental retardation and morbid obesity (for review, see Cassidy et al., 2012). The 15q11–q13 region is imprinted, leading to the expression of the SNURF-SNRPN gene and downstream noncoding region from the paternal chromosome. All paternal transcripts downstream of the *SNRPN* gene are noncoding and have been considered primarily as precursors for small RNAs (Cavaille et al., 2000; Royo and Cavaille, 2008; Runte et al., 2001). Since the minimal paternal deletion region associated with PWS (108 kb) removes only SNORD109A, the SNORD116 cluster of 29 similar snoRNAs and *IPW* (imprinted in Prader–Willi Syndrome, noncoding RNA), the most current published model is that SNORD116 deficiency is the primary cause of disease (Sahoo et al., 2008; de Smith et al., 2009; Duker et al., 2010). The function of SNORD116s is unknown. While most snoRNAs exhibit complementarity to rRNA or snRNA targets, no noncoding RNA, pre-mRNA or mRNA targets for these SNORD116 snoRNAs have been confirmed. These snoRNAs show minimal complementarity to rRNA or other RNAs so far and show greater homology to one another than to other snoRNAs. A few potential targets have been predicted bioinformatically, but none has been validated (Bazeley et al., 2008). The human and mouse SNORD116s are similar but not identical, but their deficiency in mouse only recapitulates some of the features of PWS (Skryabin et al., 2007). Thus, what function of these SNORD116 snoRNAs is and what the precise molecular cause of PWS still remains unknown.

PWS region *sno-lncRNAs* are expressed at exteremely high levels (similar in abundance to some histone mRNAs) in undifferentiated human ES cells. Interestingly, although one *sno-lncRNA* contains two snoRNAs, *sno-lncRNAs* do not colocalize with nucleoli or Cajal bodies, but rather accumulate near their sites of synthesis (Fig. 4B). Importantly, all five PWS region *sno-lncRNAs* are colocalized together in the nucleus, suggesting that they are functionally similar to each other. Importantly, although knockdown of these *sno-lncRNAs* has little effect on their parent gene, *SNURF-SNRPN*, or global gene expression, depletion of these RNAs influenced Fox2-mediated alternative splicing regulation. Each PWS region *sno-lncRNA* contains multiple binding sites for Fox2 and strongly interacts with Fox2 as demonstrated by a variety of biochemical and cellular assays (Yin et al., 2012), indicating that binding of *sno-lncRNAs* could redistribute Fox2 to specific subnuclear neighborhoods. Transcriptome analyses revealed that a number of altered exons (inclusion or exclusion of some cassette exons) were affected in opposite ways by depletion of these PWS region *sno-lncRNAs* or by knockdown of Fox2, further supporting the notion that these *sno-lncRNAs* work together to act as molecular sponges for Fox2. It is interesting to note that among the group of RNAs most affected by the depletion of PWS region *sno-lncRNAs* are several genes with a clear connection to neuronal function, thus, subtle changes of splicing regulation of these cassette exons during early embryonic development and adulthood may result in an abnormal development in PWS patients. Together, these results implicate the PWS region *sno-lncRNAs* are functionally connected to the molecular pathogenesis of PWS by acting to sequester Fox proteins. In wild type human ES cells, PWS region *sno-lncRNAs* are expressed at high levels, accumulate as strong nuclear foci near their sites of synthesis, act as molecular sponges to strongly associated with Fox2, and regulate patterns of Fox-regulated alternative splicing. While in PWS patients, all these *sno-lncRNAs* are not expressed, Fox proteins are more uniformly distributed throughout the nucleus, resulting in altered patterns of splicing regulation during early embryonic development and adulthood.

There are a number of questions that remain to be addressed for this new class of lncRNAs. First, why do PWS region *sno-lncRNAs* affect only a subset of Fox2 regulated pre-mRNA splicing? This leads to the hypothesis that pre-mRNAs with altered pattern of alternative splicing after depletion of these *sno-lncRNAs* may be located in the proximity of the PWS region. Thus, PWS region *sno-lncRNA* accumulation may offer a structural platform for such pre-mRNAs and Fox proteins to interact. Second, what is the mechanism by which PWS region *sno-lncRNAs* accumulate exclusively to their sites of synthesis after maturation? It is noteworthy that all five PWS region *sno-lncRNAs* accumulate to only one or two adjacent places near their sites of synthesis (Fig. 4B), suggesting that this localization pattern may be important for their function, such as to act as a structural platform for pre-mRNAs and proteins. Hence, answering

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

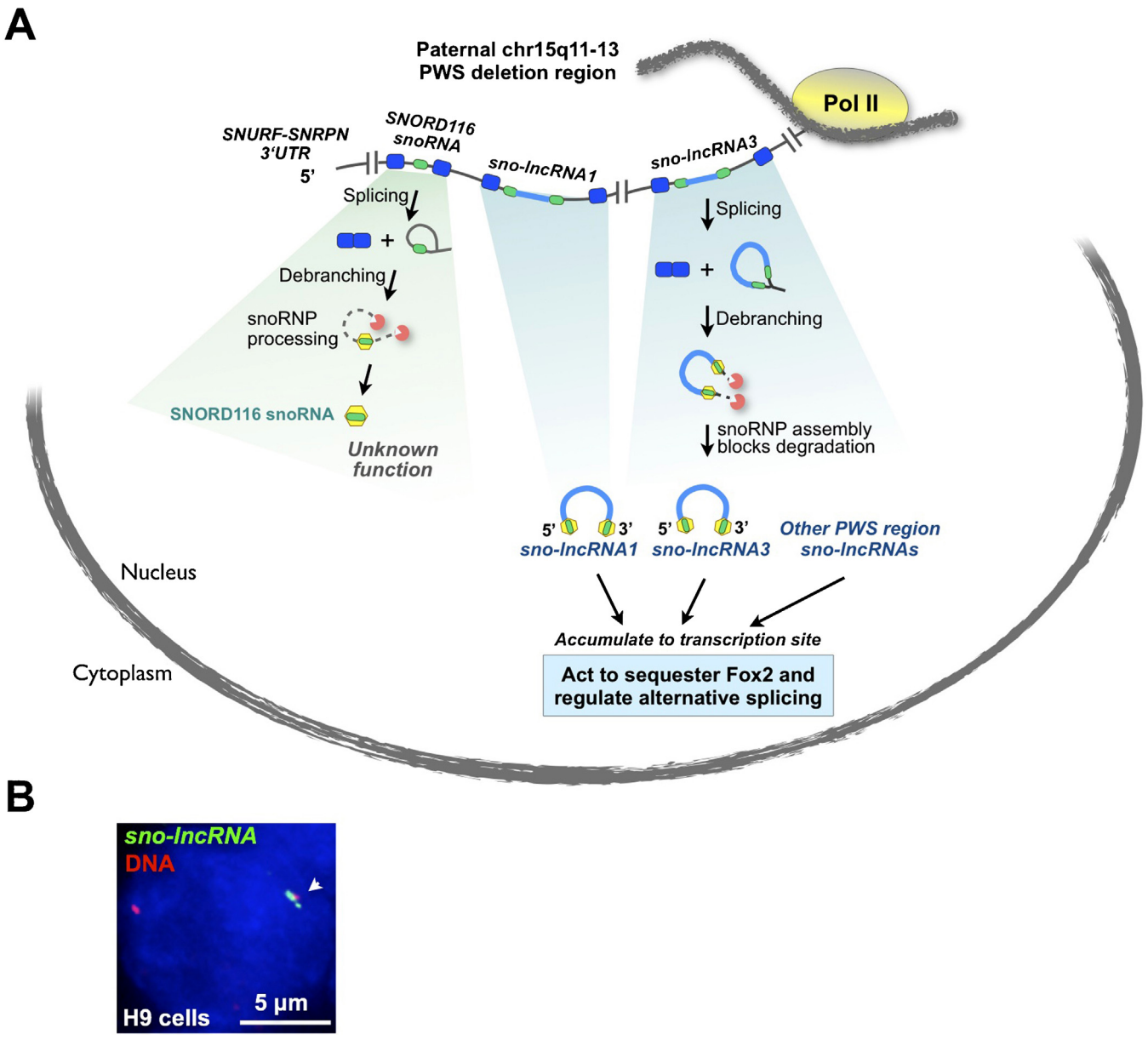

Fig. 4. *Processing and function of PWS region sno-lncRNAs.* (A) The deletion of the paternal transcribed chr15q11–13 region leads to Prader–Willi Syndrome. The minimal paternal deletion region associated with PWS (108 kb) removes snoRNAs and *sno-lncRNAs*. The processing of snoRNA is shadowed in green and the processing of *sno-lncRNA* is shadowed in blue. PWS region *sno-lncRNAs* act to sequester Fox proteins and regulate alternative splicing. See text for details. (B) All five PWS region *sno-lncRNAs* accumulate near to their sites of synthesis (a single chromosomal locus). A double FISH of *sno-lncRNAs* (green) and its adjacent DNA region (red) in H9 cells has been modified from (Yin et al., 2012). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

this question will be helpful in our understanding of whether *sno-lncRNA* nuclear accumulations represent yet another type of novel nuclear bodies. Third, is Fox2 the only protein that binds these *sno-lncRNAs*? Probably not. It is certainly possible that *sno-lncRNAs* can interact with multiple proteins to carry out functions beyond Fox protein titration. Thus, application of new approaches to systematically investigate proteins that are associated with these RNAs will be of the future interest. Fourth, how are these PWS region *sno-lncRNAs* involved in Prader–Willi pathogenesis? Although the functional characterization of these *sno-lncRNAs* revealed that they act as molecular sponges to titrate Fox proteins in undifferentiated human ES cells, their function in other cells could be different, in particular, considering the fact that their expression levels in human tissues are variable (Yin et al., 2012). Finally, despite our current understanding of *sno-lncRNAs* from PWS region, it will be of interest to learn the biogenesis and function of other type of *sno-lncRNAs*, such as the box H/ACA ended *sno-lncRNAs*. Taken together, these intron-derived and snoRNA capped lncRNAs are a class of lncRNAs with new format and what we have learned from PWS region *sno-lncRNAs* may represent only the tip of the iceberg of their regulatory functions.

4. Circular RNAs

In addition to the above mentioned linearized lncRNAs with the noncanonical 3′ ends, circular RNAs are another type of noncoding RNAs that are formed either by head-to-tail circularization from exons with covalent joined ends (Salzman et al., 2012; Jeck et al., 2013; Hansen et al., 2013; Memczak et al., 2013) or derived from the inefficient debranching of lariat introns (Zhang et al., 2013) (Fig. 1). Recent studies have demonstrated their regulatory roles in gene expression.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

4.1. Circular RNAs produced from back splice circularization

Circular RNAs were first found in the 1970s in several eukaryotic RNA viruses and were visualized to form 'pan-handle' like structures (Kolakofsky, 1976; Hsu and Coca-Prados, 1979). Several other circular RNAs produced from eukaryotic genomes have been subsequently identified, including *DCC* (*Homo sapiens* deleted in colorectal carcinoma) (Nigro et al., 1991), *c-ETS-1* (circular *ETS-1*) (Cocquerelle et al., 1992, 1993), *SRY* (sex-determining region Y) (Capel et al., 1993) and *cANRIL* (Burd et al., 2010). Although functions are largely unknown, these circular RNAs are derived from important gene loci, thus implicating that they are not junk in cells. For example, *SRY* is highly expressed in testes and plays a key role in testes development in mouse (Capel et al., 1993; Hansen et al., 2013). In addition, the *INK4a/ARF* tumor suppressor locus-associated lncRNA *ANRIL* participates directly in epigenetic transcriptional repression (Yap et al., 2010). Also, the circular form of *ANRIL* (*cANRIL*) encoded by this locus is positively correlated with *INK4/ARF* transcription (Burd et al., 2010), suggesting a role of *cANRIL* in gene expression regulation.

Very recently, the application of high-throughput sequencing technology together with computational analysis has discovered thousands of circular RNAs (circRNAs) across species, from Archaea to humans (Danan et al., 2012; Salzman et al., 2012; Jeck et al., 2013; Memczak et al., 2013). For example, by sequencing of rRNA-depleted RNAs from human fibroblasts that had been digested with RNase R - an exonuclease degrades linear but not circular RNAs, over 25,000 circular RNAs have been identified (Jeck et al., 2013). Some circular RNAs are more stable than their associated linear mRNAs (Jeck et al., 2013; Memczak et al., 2013), probabaly due to the circular structure being resistant to degradation by linear RNA decay machineries. Although their detailed biogenesis has not yet been investigated, several studies have suggested that these non-linearized RNAs are largely generated from back-spliced exons, in which splice junctions formed by an acceptor splice site at the 5′ end of an exon and a donor site at a downstream 3′ end (Burd et al., 2010; Salzman et al., 2012; Jeck et al., 2013; Hansen et al., 2013; Memczak et al., 2013) (Fig. 5). The splicing machinery is likely involved in this process although how the splicing machinery selects particular exons is still unknown. A few studies have suggested that the presence of inverted *Alu* repeats in the adjacent introns and/or exon skipping events may be involved in this process (Capel et al., 1993; Pasman et al., 1996; Jeck et al., 2013; Memczak et al., 2013).

CircRNAs largely localized to the cytoplasm but were not bound to ribosomes. Since they could be degraded by siRNAs, they were proposed to act as sponges to compete with endogenous RNAs for microRNA binding (Jeck et al., 2013). MicroRNAs are ~21-nucleotides RNAs that act by base pairing to target mRNAs and subsequently repress protein production or cause mRNA degradation (Bartel, 2004). As a microRNA can bind to many transcripts containing complementary sequences, microRNA activity can be regulated by introducing concatemers of target sites (Ebert et al., 2007), or by endogeneously expressed competing RNAs (Salmena et al., 2011). Importantly, very recent studies have demonstrated that some abundant circRNAs function as efficient "sponges" to sequester microRNAs and subsequently regulate gene expression (Hansen et al., 2013; Memczak et al., 2013). One such human circRNA, named *CDR1as* (antisense to the cerebellar degeneration-related protein 1 transcript), is predominately expressed in human and mouse brain (Hansen et al., 2011; Memczak et al., 2013). *CDR1as* is about 1.5 kb in length and contains more than 70 binding sites (among which over 60 sites are conserved) for microRNA miR-7

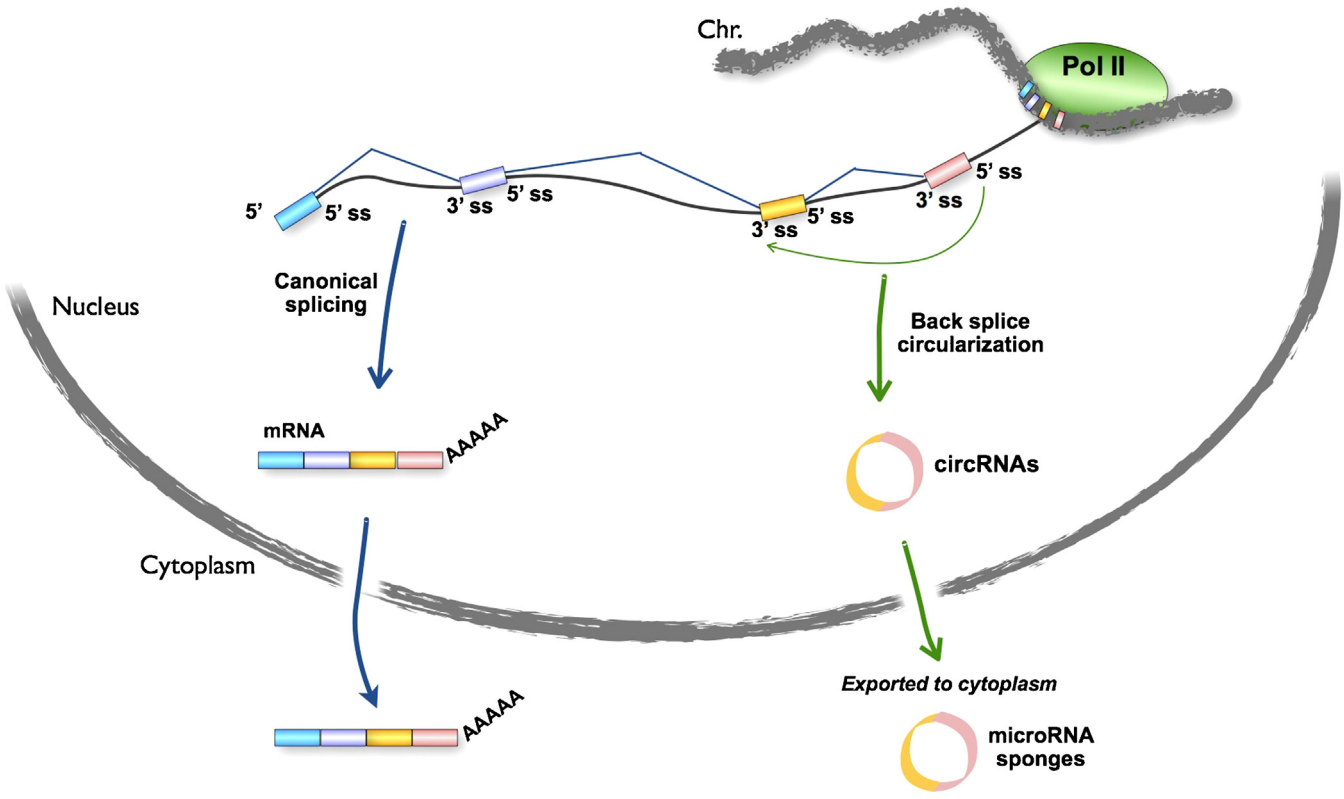

Fig. 5. *CircRNAs produced from back-splice circularization function as microRNA sponges.* Canonical splicing occurs after or concurrently with transcription in which introns are removed and exons are joined, and the mature 3′ end is generated by endonucleolytic cleavage followed by the addition of a poly(A) tail in a non-templated fashion. CircRNAs are processed from back-spliced exons, in which splice junctions formed by an acceptor splice site at the 5′ end of an exon and a donor site at a downstream 3′ end. CircRNAs are largely localized to the cytoplasm and some can function as microRNA sponges.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

(Memczak et al., 2013; Hansen et al., 2013), thus, *CDR1as* was also named *ciRS-7* (circular RNA sponge for miR-7) (Hansen et al., 2013). This highly abundant circRNA (about the 1/5 abundance to *GAPDH* mRNAs) has the potential to bind up to ~20,000 miR-7 molecules per cell (Memczak et al., 2013), thereby preventing miR-7 from binding other mRNAs. Consistant with the sponge model, this circular RNA was found to be associated with the miRNA effector AGO proteins in a miR-7 dependent manner and decreased expression of this circRNA caused the reduced expression of mRNAs containing miR-7 binding sites (Hansen et al., 2013; Memczak et al., 2013).

Thus, a role as a molecular sponge for microRNAs classifies circRNAs in a similar category with several other recently reported competing endogenous RNAs and pseudogene RNAs (Cesana et al., 2011; Tay et al., 2011; Salmena et al., 2011). It is of great interest to investigate the participation of thousands of other circRNAs in microRNA binding as molecular sponges. In addition, several linear lncRNAs have been recently characterized as molecular sponges for RNA binding proteins (Tollervey et al., 2011; Polymenidou et al., 2011; Yin et al., 2012), which further indicate that circRNAs may act to sequester RNA-binding proteins beyond their roles in regulating microRNAs. Similarly, circRNAs may even base pair with other larger cellular RNAs, and together with associated proteins, resulting in the formation of (functional) RNA-protein complexes in the cytoplasm.

4.2. Circular intronic RNAs derived from lariat introns

Very recently, we have identified another type of circular RNAs, circular intronic RNAs (ciRNAs), derived from lariat introns (Zhang et al., 2013). Intron lariats are short-lived intermediates during RNA splicing. An intron lariat carrying a 2′,5′-phosphodiester bond is debranched and rapidly degraded by exonuclease trimming during splicing, thus usually rendering these molecules undetectable by biochemical assays. However, a few exceptions were noted in earlier studies. For example, a spliced intron was found to accumulate as a set of lariat RNA structures with different length of tails in the nucleus, resulting in stable intermediates of intron lariats (Qian et al., 1992). In addition, stable lariat introns of some viruses were also seen in infected cells due to atypical splicing (Wu et al., 1998).

By deep sequencing of RNAs collected from a combination treatments of poly(A)- selection, rRNA depletion and RNaseR digestion followed by new computational algorithm analyses, hundreds of ciRNAs were suprisingly identified in human cells (Zhang et al., 2013). These stable ciRNAs are derived from intron lariats, escape from debranching and depend on a consensus RNA motif containing a 7 nt GU rich element near 5′ splice site and an 11 nt GU rich element branch point (Fig. 6). Although the resulting ciRNA carries a 2′,5′-phosphodiester bond to form the circular structure, it does not contain the 3′ linear components (the sequence from the 3′ end of an intron to its branch point site) (Zhang et al., 2013), suggesting that they are different from splicing intermediates but rather represent a new class of circular RNAs derived from excised introns. However, how these consensus RNA elements work to resist debranching and what protein partners are involved in this process still remain to be determined.

Some abundant ciRNAs are much more stable than their associated parent linear mRNAs, likely due to their circular structure protection. Strikingly, different from circRNAs produced from back splice circularization (Jeck et al., 2013; Hansen et al., 2013; Memczak et al., 2013) (Fig. 5), these intron-derived ciRNAs have little enrichment for microRNA target sites, suggesting that they are fucntionally distinct. Detailed studies have revealed that some abundant expressed ciRNAs accumulate in the nucleus and are largely associated with their sites of transcription, where they specifically interact with the elongating Pol II complex and act as positive regulator for Pol II transcription to regulate local gene

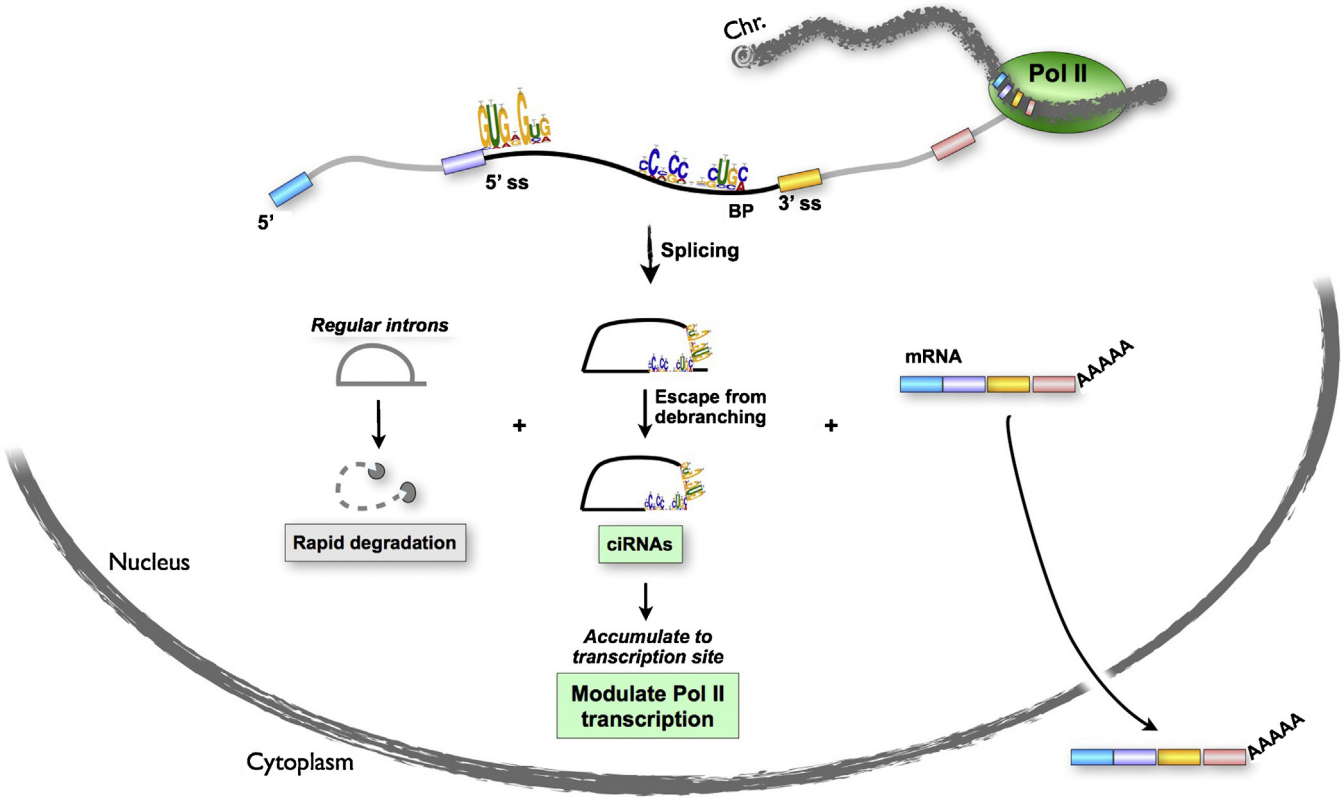

Fig. 6. *CiRNAs derived from lariat introns act as Pol II regulators*. Introns are usually degraded after splicing. Introns carrying consensus motifs near the 5′ splice site and the branch point site lead to a failure in debranching of lariat introns and the formation of ciRNAs with a 2′,5′-phosphodiester bond. CiRNAs are abundant in the nucleus and some can act as positive regulators of Pol II transcription.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

expression (Zhang et al., 2013). Interestingly, although ciRNAs are widely expressed in human cells and tissue samples, they are less evolutionarily conserved (Zhang et al., 2013). Comparison of human ciRNA producing introns with those in the mouse genome revealed that the consensus RNA motif that is key for ciRNA formation is not conserved in mouse, suggesting that evolutionarily unconserved intronic noncoding transcripts could add more complexity to human genomes.

5. Perspectives

Although the great majority of eukaryotic RNAs end in a poly(A) tail, recent advances in deep sequencing and computational analysis have uncovered thousands of previously unknown RNAs with a variety of new processing formats. This has not only provided new insights into the "dark matter" of human genome, but also raised a number of questions and issues to be addressed. First, as the expression of non-canonical lncRNAs is tissue-/cell- specific and evolutionarily less conserved, further application of similar approaches for different tissues and species may result in the identification of additional new RNAs. Second, beyond the above-mentioned new types of RNAs, what other RNAs might have been missed from the current lncRNA study? The combination of different treatments to RNAs prior to library preparation for RNA sequencing may allow one to uncover new RNA species. Third, it is noteworthy that the above-mentioned RNAs with non-canonical 3′ ends are more stable than linear RNAs, suggesting that their cellular fate could be different from RNAs with canonical poly(A) tails, probably due to the non-canonical end or circular structure being able to protect them from linear RNA decay machinery. However, such RNAs do have a finite lifetime. How are they processed for degradation? A recent study revealed that Drosha-DGCR8 microprocessor, which is required for microRNA biogenesis, is also involved in controlling the abundance of *MALAT1* (Macias et al., 2012). Thus, it will be of great interest to learn what other RNase enzymes or cellular decay machineries are involved in the processing of such RNAs. Finally, it is known that the coupling between transcription and processing is key for the functionality of some RNAs. Although the biogenesis of these new RNAs has just begun to be appreciated, what we have learned from their processing is already distinct from that of canonical RNAs. As recent studies have suggested that these new RNAs can participate in diverse aspects of biological regulation by a variety of mechanisms, understanding the coupling between transcription and processing of these new lncRNAs will provide additional insights into their biological roles.

Acknowledgements

We are grateful to G. Carmichael for critical reading of the manuscript. This work was supported by grants XDA01010206 and 2012OHTP08 from CAS, grants 2014CB964800 and 2011CBA01105 from MOST, grant 31271376 from NSFC, grants 2012SSTP01 and 2012KIP101 from SIBS.

References

Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.
Bazeley PS, Shepelev V, Talebizadeh Z, Butler MG, Fedorova L, Filatov V, et al. snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. Gene 2008;408:172–9.
Bernard D, Prasanth KV, Tripathi V, Colasse S, Nakamura T, Xuan Z, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J 2010;29:3082–93.
Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. The multifunctional nucleolus. Nat Rev Mol Cell Biol 2007;8:574–85.
Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol 2009;186:637–44.
Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENbeta noncoding RNAs. Proc Nat Acad Sci USA 2012;109:19202–7.
Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. PLoS Genet 2010;6:e1001233.
Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011;25:1915–27.
Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. Cell 1993;73:1019–30.
Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader–Willi syndrome. Genet Med 2012;14:10–26.
Cavaille J, Buiting K, Kiefmann M, Lalande M, Brannan CI, Horsthemke B, et al. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc Nat Acad Sci USA 2000;97:14311–6.
Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 2011;147:358–69.
Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol Cell 2009;35:467–78.
Chen LL, DeCerbo JN, Carmichael GG. Alu element-mediated gene silencing. EMBO J 2008;27:1694–705.
Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science 2005;308:1149–54.
Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, et al. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell 2009;33:717–26.
Cocquerelle C, Daubersies P, Majerus MA, Kerckaert JP, Bailleul B. Splicing with inverted order of exons occurs proximal to large introns. EMBO J 1992;11:1095–8.
Cocquerelle C, Mascrez B, Hetuin D, Bailleul B. Mis-splicing yields circular RNA molecules. FASEB J 1993;7:155–60.
Danan M, Schwartz S, Edelheit S, Sorek R. Transcriptome-wide discovery of circular RNAs in Archaea. Nucleic Acids Res 2012;40:3131–42.
De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, et al. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 2010;8:e1000384.
de Smith AJ, Purmann C, Walters RG, Ellis RJ, Holder SE, Van Haelst MM, et al. A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. Hum Mol Genet 2009;18:3257–65.
Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012;22:1775–89.
Duker AL, Ballif BC, Bawle EV, Person RE, Mahadevan S, Alliman S, Thompson R, et al. Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116C/D box snoRNA cluster in Prader–Willi syndrome. Eur J Hum Genet 2010;18:1196–201.
Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 2007;4:721–6.
Eissmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, Gross M, et al. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol 2012;9:1076–87.
Filipowicz W, Pogacic V. Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 2002;14:319–27.
Fox AH, Bond CS, Lamond AI. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. Mol Biol Cell 2005;16:5304–15.
Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, et al. Paraspeckles: a novel nuclear domain. Curr Biol 2002;12:13–25.
Fox AH, Lamond AI. Paraspeckles. Cold Spring Harbor Perspect Biol 2010;2:a000687.
Gu H, Das Gupta J, Schoenberg DR. The poly(A)-limiting element is a conserved *cis*-acting sequence that regulates poly(A) tail length on nuclear pre-mRNAs. Proc Nat Acad Sci USA 1999;96:8943–8.
Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 1983;35:849–57.
Gutschner T, Baas M, Diederichs S. Noncoding RNA gene silencing through genomic integration of RNA destabilizing elements using zinc finger nucleases. Genome Res 2011;21:1944–54.
Gutschner T, Hammerle M, Diederichs S. MALAT1-a paradigm for long noncoding RNA function in cancer. J Mol Med (Berlin) 2013a;91:791–801.
Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, Hung G, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res 2013b;73:1180–9.
Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223–7.
Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 2011;477:295–300.
Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature 2012;482:339–46.
Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. Natural RNA circles function as efficient microRNA sponges. Nature 2013;495:384–8.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. EMBO J 2011;30:4414–22.
Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, et al. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. Nat Methods 2013;10:133–9.
Hsu MT, Coca-Prados M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. Nature 1979;280:339–40.
Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics 2007;8:39.
Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA 2013;19:141–57.
Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene 2003;22:8031–41.
Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. Antisense transcription in the mammalian transcriptome. Science 2005;309:1564–6.
Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Nat Acad Sci USA 2009;106:11667–72.
Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature 2010;465:182–7.
Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J 2001;20:3617–22.
Kiss T, Filipowicz W. Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. Genes Dev 1995;9:1411–24.
Kolakofsky D. Isolation and characterization of Sendai virus DI-RNAs. Cell 1976;8:547–55.
Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 2013;498:516–20.
Lin R, Maeda S, Liu C, Karin M, Edgington TS. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. Oncogene 2007;26:851–8.
Macias S, Plass M, Stajuda A, Michlewski G, Eyras E, Caceres JF. DGCR8 HITS-CLIP reveals novel functions for the microprocessor. Nat Struct Mol Biol 2012;19:760–6.
Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. Nature 2006;444:953–6.
Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol 2011a;13:95–101.
Mao YS, Zhang B, Spector DL. Biogenesis and funciton of nuclear bodies. Trends Genet 2011b;27:295–306.
Marvin MC, Clauder-Munster S, Walker SC, Sarkeshik A, Yates JR, Steinmetz LM. Accumulation of noncoding RNA due to an RNase P defect in *Saccharomyces cerevisiae*. RNA 2011;17:1441–50.
Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol 2007;8:209–20.
Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 2013;495:333–8.
Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science 2010;330:1244–7.
Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. EMBO J 2012;31:4020–34.
Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, et al. Malat1 is not an essential component of nuclear speckles in mice. RNA 2012;18:1487–99.
Nakagawa S, Naganuma T, Shioi G, Hirose T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. J Cell Biol 2011;193:31–9.
Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, et al. Scrambled exons. Cell 1991;64:607–13.
Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 2002;420:563–73.
Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010;143:46–58.
Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 2008;40:1413–5.
Pasman Z, Been MD, Garcia-Blanco MA. Exon circularization in mammalian nuclear extracts. RNA 1996;2:603–10.
Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci 2011;14:459–68.
Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, et al. Regulating gene expression through RNA nuclear retention. Cell 2005;123:249–63.
Qian L, Vu MN, Carter M, Wilkinson MF. A spliced intron accumulates as a lariat in the nucleus of T cells. Nucleic Acids Res 1992;20:5345–50.
Reiner R, Alfiya-Mor N, Berrebi-Demma M, Wesolowski D, Altman S, Jarrous N. RNA binding properties of conserved protein subunits of human RNase P. Nucleic Acids Res 2011;39:5704–14.
Richard P, Manley JL. Transcription termination by nuclear RNA polymerases. Genes Dev 2009;23:1247–69.
Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012;81:145–66.
Rodriguez-Trelles F, Tarrio R, Ayala FJ. Origins and evolution of spliceosomal introns. Annu Rev Genet 2006;40:47–76.
Royo H, Cavaille J. Non-coding RNAs in imprinted gene clusters. Biol Cell 2008;100:149–66.
Runte M, Huttenhofer A, Gross S, Kiefmann M, Horsthemke B, Buiting K. The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum Mol Genet 2001;10:2687–700.
Saha S, Murthy S, Rangarajan PN. Identification and characterization of a virus-inducible non-coding RNA in mouse brain. J Gen Virol 2006;87:1991–5.
Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, et al. Prader–Willi phenotype caused by paternal deficiency for the HBII-85C/D box small nucleolar RNA cluster. Nat Genet 2008;40:719–21.
Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language. Cell 2011;146:353–8.
Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS One 2012;7:e30733.
Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proc Nat Acad Sci USA 2009;106:2525–30.
Shatkin AJ, Manley JL. The ends of the affair: capping and polyadenylation. Nat Struct Biol 2000;7:838–42.
Skryabin BV, Gubar LV, Seeger B, Pfeiffer J, Handel S, Robeck T, et al. Deletion of the MBII-85 snoRNA gene cluster in mice results in postnatal growth retardation. PLoS Genet 2007;3:e235.
Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. J Cell Sci 2007;120:2498–506.
Souquere S, Beauclair G, Harper F, Fox A, Pierron G. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. Mol Biol Cell 2010;21:4020–7.
Spector DL, Lamond AI. Nuclear speckles. Cold Spring Harbor Perspect Biol 2011;3.
Sun L, Campbell FE, Yandek LE, Harris ME. Binding of C5 protein to P RNA enhances the rate constant for catalysis for P RNA processing of pre-tRNAs lacking a consensus (+ 1)/C(+ 72) pair. J Mol Biol 2010;395:1019–37.
Sun L, Campbell FE, Zahler NH, Harris ME. Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P. EMBO J 2006;25:3998–4007.
Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res 2009;19:347–59.
Tay Y, Kats L, Salmena L, Weiss D, Tan SM, Ala U, et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell 2011;147:344–57.
Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 2005;33:201–12.
Tian B, Manley JL. Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem Sci 2013;38:312–20.
Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat Neurosci 2011;14:452–8.
Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 2010;39:925–38.
Tripathi V, Shen Z, Chakraborty A, Giri S, Freier SM, Wu X, et al. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. PLoS Genet 2013;9:e1003368.
Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA. Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. Cell Rep 2012;2:26–32.
Tycowski KT, Shu MD, Steitz JA. A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Genes Dev 1993;7:1176–90.
Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell 2013;154:26–46.
Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 2011;147:1537–50.
Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, et al. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 2008;453:1239–43.
Wilusz JE, Freier SM, Spector DL. 3′ End processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell 2008;135:919–32.
Wilusz JE, Whipple JM, Phizicky EM, Sharp PA. tRNAs marked with CCACCA are targeted for degradation. Science 2011;334:817–21.
Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. Genes Dev 2012;26:2392–407.
Wu Q, Kim YC, Lu J, Xuan Z, Chen J, Zheng Y, et al. Poly A-transcripts expressed in HeLa cells. PLoS One 2008;3:e2803.
Wu TT, Su YH, Block TM, Taylor JM. Atypical splicing of the latency-associated transcripts of herpes simplex type 1. Virology 1998;243:140–9.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009

Yang L, Altman S. A noncoding RNA in *Saccharomyces cerevisiae* is an RNase P substrate. RNA 2007;13:682–90.
Yang L, Duff MO, Graveley BR, Carmichael GG, Chen LL. Genomewide characterization of non-polyadenylated RNAs. Genome Biol 2011a;12:R16.
Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, et al. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell 2011b;147:773–88.
Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 2010;38:662–74.
Yin QF, Yang L, Zhang Y, Xiang JF, Wu YW, Carmichael GG, et al. Long noncoding RNAs with snoRNA ends. Mol Cell 2012;48:219–30.
Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a *cis*-regulatory role in the adult. Cell Rep 2012;2:111–23.
Zhang Y, Zhang XO, Tian C, Xiang JF, Yin QF, Xing YH, et al. Circular intronic long noncoding RNAs. Mol Cell 2013;51:792–806.

Please cite this article in press as: Zhang Y, et al. Life without A tail: New formats of long noncoding RNAs. Int J Biochem Cell Biol (2014), http://dx.doi.org/10.1016/j.biocel.2013.10.009