Trends in Biochemical Sciences

CellPress

Series: Fresh Perspectives from Emerging Experts

Review

Linking Long Noncoding RNA Localization and Function

Ling-Ling Chen[1,2,*]

Recent studies have revealed the regulatory potential of many long noncoding RNAs (lncRNAs). Most lncRNAs, like mRNAs, are transcribed by RNA polymerase II and are capped, polyadenylated, and spliced. However, the subcellular fates of lncRNAs are distinct and the mechanisms of action are diverse. Investigating the mechanisms that determine the subcellular fate of lncRNAs has the potential to provide new insights into their biogenesis and specialized functions.

The Diversity of Long Noncoding RNAs

By their broadest definition, long noncoding RNAs (lncRNAs) comprise different types of RNA polymerase II (Pol II)-transcribed molecules with sizes greater than 200 nt in length, which distinguishes them from short RNAs such as microRNAs (Figure 1). It is now widely accepted that lncRNAs have regulatory roles in gene expression at both the transcriptional and post-transcriptional levels in diverse cellular contexts and biological processes. lncRNAs are responsible for nuclear structure integrity, and can regulate the expression of either nearby genes (acting in *cis* in the nucleus) or genes elsewhere in cells (acting in *trans* in the nucleus or cytoplasm) by interacting with proteins, RNAs, and DNAs. lncRNAs can positively or negatively regulate gene expression through multiple mechanisms, including through recruiting transcription factors or chromatin-modifying complexes to their DNA targets, by forming heterogeneous nuclear ribonucleoprotein (hnRNP) complexes, by acting as decoys to sequester RNA-binding proteins (RBPs) and microRNAs, or by directly interacting with RNAs and DNAs by base pairing (see reviews [1–3]).

Most annotated lncRNAs are Pol II transcribed, hence they are presumably capped, polyadenylated, and spliced just like mRNAs [4,5]. Such lncRNAs are transcribed either from intergenic regions of the genome [i.e., large intergenic noncoding RNAs (lincRNAs); Figure 1A] or from the opposite strand of protein-coding genes [natural antisense transcripts (NATs); Figure 1B] [6]. lincRNAs are the most abundant class of lncRNAs comprising over 10 000 species so far. A major difference between these lncRNAs and mRNAs is that lncRNAs have little if any protein-coding potential. lncRNAs may contain fewer exons than mRNAs and often have weak cryptic splicing and polyadenylation signals [5].

While a large proportion of lncRNAs look like mRNA transcripts, a number of lncRNAs form their 3′ end in unusual ways (Box 1 and Figures 1C–F) and others are processed from long polyadenylated primary transcripts (Box 1 and Figures 1G–I). In addition, new types of lncRNAs are likely to be discovered through integrated approaches. Recently, the advent of RNA-centric methods combined with deep sequencing as well as emerging genome-editing technologies have greatly advanced our knowledge of both the functional significance and mechanisms of action of lncRNAs (see reviews [3,7]). Despite these rapid areas of growth, we still know little about the life cycle of most 'mRNA-like' lncRNAs. How their fates and functions are different from each other and how cells distinguish lncRNAs from mRNAs remain obscure. Importantly, it is

Trends

Long noncoding RNAs (lncRNAs) comprise different classes of RNA molecules with sizes greater than 200 nt.

The function of lncRNAs is associated with their unique subcellular localization patterns.

Mature lncRNAs can accumulate in *cis*, localize in the nucleus in *trans*, or export to the cytoplasm to execute their functions.

Multiple factors including ribonucleic nuclear retention elements, nuclear protein factors, higher-order chromosome organization, and the coupling of RNA–protein assemblies with lncRNA transcription may dictate the subcellular localization patterns of lncRNAs.

A better understanding of the lncRNAs themselves is crucial to link these noncoding transcripts to RNA biology and to address their cellular roles in depth.

[1]State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Key Laboratory of Molecular Andrology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China
[2]School of Life Science and Technology, Shanghai Tech University, 100 Haike Road, Shanghai, China

*Correspondence:
linglingchen@sibcb.ac.cn (L.-L. Chen).

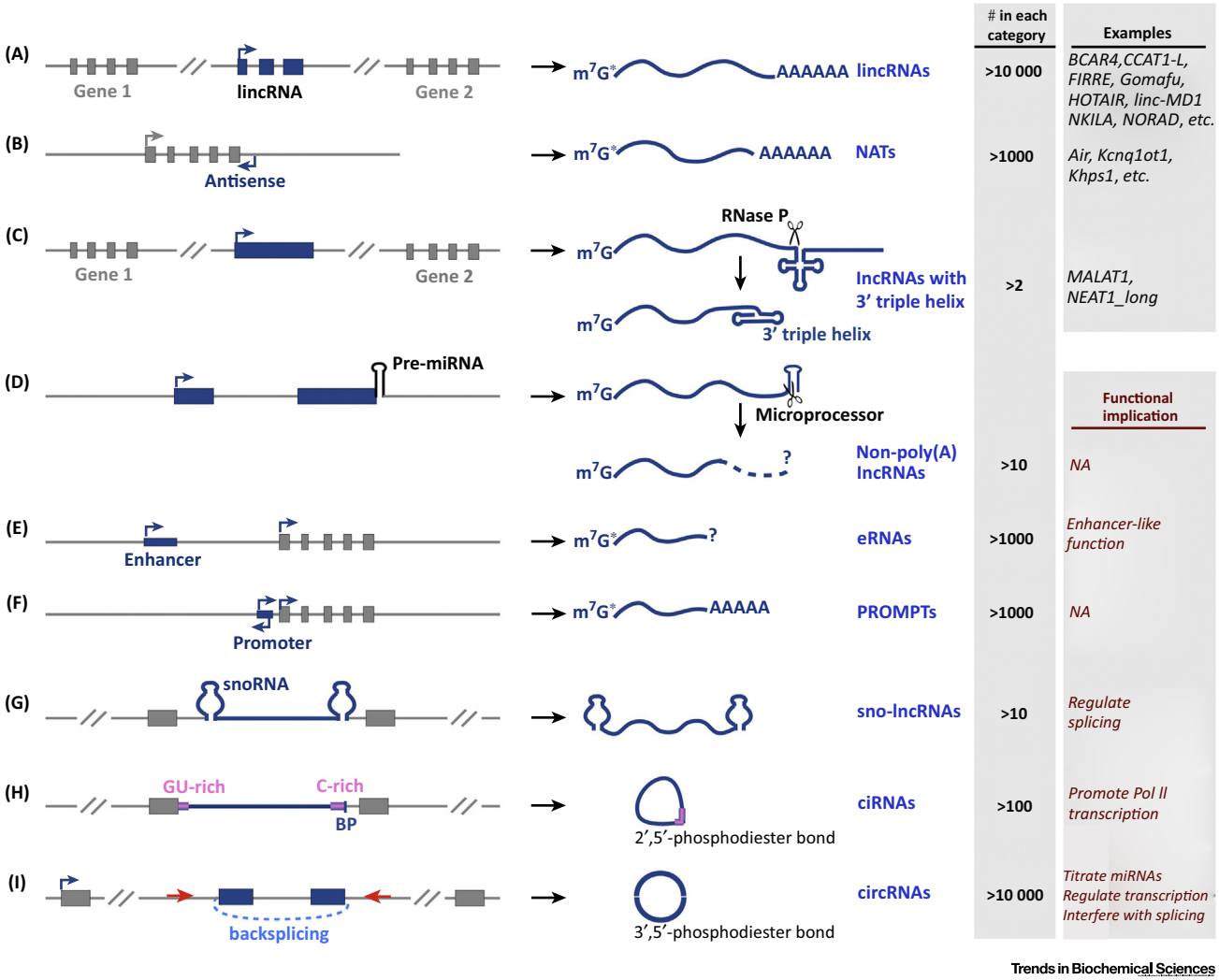

Figure 1. The Diversity of Long Noncoding RNAs (lncRNAs) in Mammalian Genomes. (A) Large intergenic noncoding RNAs (lincRNAs). (B) Natural antisense transcripts (NATs). (C and D) lncRNAs with alternatively processed 3′ ends by ribonuclease P (RNase P; C) or microprocessor (D) cleavage. (E) Enhancer RNAs (eRNAs). (F) Promoter upstream transcripts (PROMPTs). (G) Small nucleolar RNA (snoRNA)-ended lncRNAs (sno-lncRNAs). (H) Circular intronic RNAs (ciRNAs). ciRNA formation depends on consensus RNA sequences (pink bars) to avoid debranching of the lariat intron. (I) Circular RNAs (circRNAs) produced from backsplicing of exons. Inverted complementary sequences (red arrows) in introns flanking circularized exons promote circRNA biogenesis. Gray columns: left, the number of lncRNAs identified in each category in mammals so far; right, examples and functional implication of lncRNAs of each category. *These Pol II transcripts are presumably capped, but the direct lines of evidence showing the presence of m^7G cap at their 5′ termini are still warranted. Abbreviations: BCAR4, Breast cancer antiestrogen resistance 4; BP, branch point; CCAT1-L, colon cancer associated transcript 1-long isoform; FIRRE, functional intergenic repeating RNA element; HOTAIR, HOX transcript antisense RNA; NKILA, nuclear factor-κB interacting lncRNA; NORAD, noncoding RNA activated by DNA damage; Pol II, polymerase II.

becoming increasingly clear that the function of lncRNAs is associated with their unique subcellular localization, but how cells sort different lncRNAs to specific subcellular compartments for carrying out their roles in gene regulation remains unclear. Understanding the mechanisms of how the subcellular fate of lncRNAs is achieved and regulated has the potential to provide new insights into their biogenesis and functions.

The Function of lncRNAs Are Associated with Their Subcellular Fates

Like proteins, the function of lncRNAs depends on their subcellular localization. Many lncRNAs are recognized as important modulators for nuclear functions (see reviews [8,9]) and exhibit distinct nuclear localization patterns (Figures 2A–D). Others must be exported to the cytoplasm

Box 1. Processing of Long Noncoding RNAs (lncRNAs) in Alternative Ways.

First, MALAT1 and NEAT1_long are two lncRNAs that are localized to nuclear speckles [68] and paraspeckles [31], respectively. They are processed at their 3′ ends by recognition and cleavage of tRNA-like structures by ribonuclease P (which processes the 5′ ends of tRNAs). Ribonuclease P cleavage leads to the formation of their mature 3′ ends, which are subsequently protected by a conserved stable U-A•U triple-helical RNA structure ('•' denotes the Hoogsteen face and '-' denotes the Watson–Crick face) [60,61] (see Figure 1C in main text). Second, some lncRNA transcripts containing miRNAs (lnc-pri-miRNAs) use microprocessor cleavage to terminate transcription, resulting in unstable lncRNAs without 3′ end poly(A) tails [69] (see Figure 1D in main text). Third, enhancer RNAs (eRNAs) are bidirectionally transcribed from enhancers by polymerase II (Pol II) and lack poly(A) tails (see Figure 1E in main text) [70,71] but have enhancer-like functions [72–74]. Depletion of Integrator, a complex with an RNA endonuclease activity, resulted in the accumulation of primary eRNA transcripts that remained bound to Pol II, suggesting a role of Integrator in cleaving the 3′ end of eRNAs [75]. Fourth, promoter upstream transcripts (PROMPTs) are bidirectionally transcribed from promoters of protein-coding genes. They are 200–600 nt in length, capped, and polyadenylated (see Figure 1F in main text). They are nuclear retained and undergo rapid degradation [76,77]. PROMPTs may not have a function. However, the rapid degradation of transcribed PROMPTs may relate to the choice of promoter directionality [78]. Both PROMPTs and eRNAs appear to be degraded by the RNA exosome [79]. Fifth, excised introns can also generate lncRNAs. Small nucleolar RNA-ended lncRNAs (sno-lncRNAs) are formed when one intron contains two snoRNA genes (see Figure 1G in main text). During splicing, the sequences between snoRNAs are not degraded, leading to the accumulation of lncRNAs flanked by snoRNAs but lacking 5′ caps and 3′ poly(A) tails in the nucleus [20]. Their expression is species specific and results from species-specific alternative splicing [80]. Finally, circular RNAs are yet another type of lncRNAs that are protected from RNA exonucleolytic decay machineries. They can be derived from the inefficient debranching of lariat introns (see Figure 1H in main text) [63] or from pre-mRNAs backsplicing of exons (see Figure 1I in main text). Circular RNAs may regulate gene expression by titrating miRNAs, regulating transcription, or interfering with splicing (see review [81]).

to carry out their regulatory roles (Figure 2E). In this review, a number of well-characterized lncRNAs are classified into three groups depending on their subcellular localization to illustrate the association of lncRNA localization and function: those that are absolutely nuclear localized in *cis* (Figures 2A,B), those that are mainly nuclear localized and function in *trans* (Figures 2C,D), and those that largely localize and function in the cytoplasm (Figure 2E). It is worth noting that a recent large-scale evaluation of the subcellular fates of lncRNAs in human cell lines using single-molecule RNA fluorescence *in situ* hybridization revealed that lncRNAs exhibited a wide range of subcellular localization patterns, including not only distinct patterns of nuclear localization but also nonspecific location in both the nucleus and cytoplasm [10].

lncRNAs That Accumulate in *cis*

Of those that have nuclear functions, one group of lncRNAs accumulates at their sites of transcription (in *cis*; Figures 2A,B). These lncRNAs can be classified into two subgroups based on whether they act in *cis*, close to the site of transcription, or in *trans*. The former lncRNAs execute their regulatory roles in *cis* by various mechanisms (Figure 2A). For example, they can modulate local gene expression either by recruiting transcription factors, chromatin organizers, or chromatin modifiers, or by forming a DNA–RNA triplex that anchors the lncRNA and associated effector proteins to the gene promoter. For example, a set of lncRNAs is induced by DNA damage from the promoter of cyclin D1 gene (*CCND1*); these lncRNAs bind to the translocated in liposarcoma (TLS) protein and allosterically modulate TLS activity to inhibit histone acetyltransferase in *cis*, which suppresses *CCND1* transcription [11]. Furthermore, regulatory RNAs transcribed from promoter-proximal and promoter-distal enhancers remain bound to chromatin and contribute to stable transcription factor (such as YY1) occupancy at these sites to maintain gene-expression programs [12]. In addition, the lncRNA CCAT1-L is transcribed from the upstream super-enhancer region of the *MYC* gene; CCAT1-L accumulates in *cis*, interacts with the chromatin organizer CTCF, and promotes chromosome looping between the super-enhancer and the *MYC* promoter, resulting in enhanced *MYC* transcription in colorectal cancer [13].

Several well-characterized NATs accumulate to their sites of transcription and affect expression of neighboring genes. In mammalian imprinting, Air and kcnq1ot1 lncRNAs are involved in the allele-specific silencing of their *cis*-linked genes [14–16]. Air recruits H3K9 histone methyltransferase G9a to maintain the imprinted silencing of multiple genes including the Igf2r locus [14].

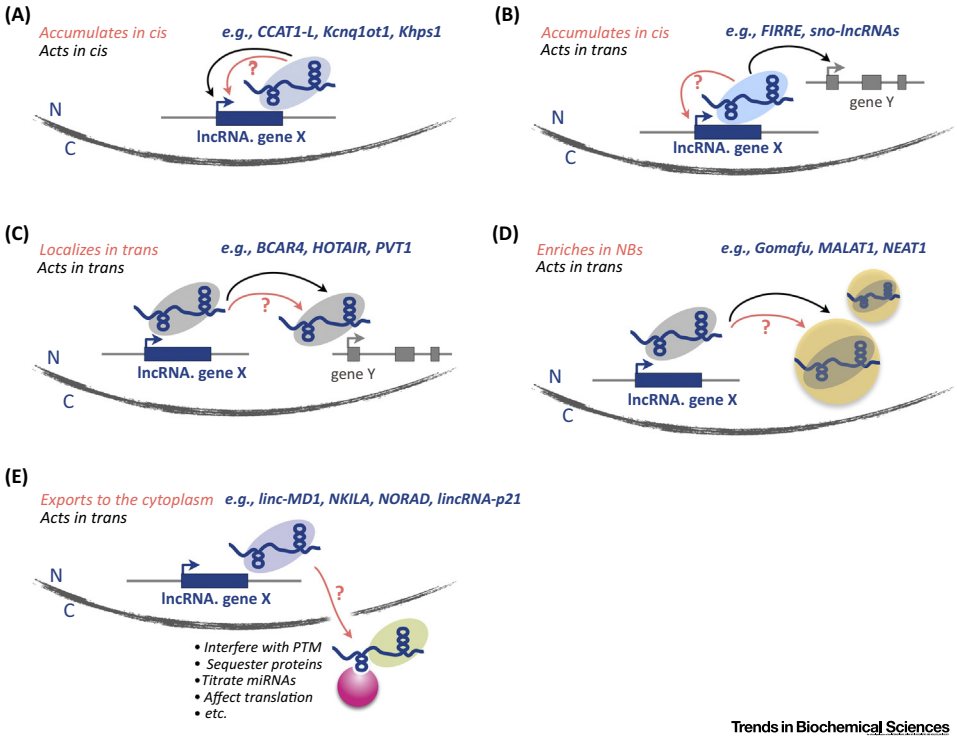

Figure 2. Functions of Long Noncoding RNAs (lncRNAs) Are Associated with Their Subcellular Fates. lncRNAs have distinct subcellular localization patterns, allowing lncRNAs to execute their specified functions. (A) lncRNAs can accumulate and act in *cis* once they are transcribed. (B) lncRNAs can accumulate in *cis* once they are transcribed, but act in *trans* that affects genes located in the same chromosome at a distance or in different chromosomes. (A and B) Mechanisms of lncRNA in *cis* localization are largely unknown. (C) lncRNAs can localize to elsewhere in the nucleoplasm in *trans* and act in *trans*. (D) lncRNAs can accumulate to specific nuclear bodies (orange circles) and act in *trans*. The mechanisms of lncRNA nuclear retention remain largely unknown, as does whether such lncRNAs are required to be translocated in the nucleus to regulate gene expression. (E) lncRNAs can be exported to the cytoplasm to execute their functions. For example, a cytoplasmic lncRNA can sequester protein (pink circle) or interfere with protein post-translational modification (PTM). Whether the nucleocytoplasmic export of cytoplasmic lncRNAs is distinct from that of mRNA has not yet been examined. The color of shaded oval indicates differences in protein composition of long noncoding ribonucleoproteins. Pink arrows, the unknown mechanisms for lncRNAs to gain specific subcellular localization patterns; black arrows, lncRNAs execute functions in distinct subcellular compartments. BCAR4, breast cancer antiestrogen resistance 4; CCAT1-L, colon cancer associated transcript 1-long isoform; FIRRE, functional intergenic repeating RNA element; lincRNA, large intergenic noncoding RNAs; MALAT1, metastasis-associated lung adenocarcinoma transcript 1; NBs, nuclear bodies; NEAT1, nuclear-enriched abundant transcript 1; NKILA, nuclear factor-κB interacting lncRNA; NORAD, noncoding RNA activated by DNA damage; PVT1, plasmacytoma variant translocation 1; Sno-lncRNA, small nucleolar RNA-ended lncRNAs.

By contrast, kcnq1ot1 recruits G9a and the H3K27 histone methyltransferase PRC2 (polycomb repressive complex 2), which leads to the bidirectional silencing of a dozen paternally imprinted genes in the *Kcnq1* locus [15,16]. Finally, the antisense lncRNA Khps1 activates proto-oncogene *SPHK* transcription in an E2F1-dependent manner. This is achieved by forming a DNA–RNA triplex of the *SPHK1* promoter with Khps1, which further recruits the histone acetyltransferase p300/CBP to the *SPHK1* promoter, leading to changes in chromatin structure and enhancing E2F1-dependent expression of SPHK1 [17]. Importantly, the act of transcription itself may influence adjacent gene expression in addition to the effects of the in-*cis*-accumulated lncRNA products. For instance, the imprinted *Igf2r* silencing also requires Air transcriptional overlap of the *Igf2r* promoter, which interferes with Pol II recruitment in the absence of repressive chromatin [18].

In-*cis*-localized lncRNAs can also influence gene expression in *trans* (Figure 2B). The lncRNA FIRRE is transcribed from the X chromosome and involved in the formation of five *trans*-chromosomal contacts at its site of transcription [19]. This spatial colocalization is achieved

through the 3D organization of the chromosomes, which enables long-range interactions between the regulatory genomic elements and the target gene locus. Such colocalization patterns are dramatically decreased in cells that lack the *Firre* locus [19]. In addition, there is yet another mechanism through which in-*cis*-accumulated lncRNAs can act as decoys for RBPs to achieve gene regulation in *trans*. For example, in human embryonic stem cells, small nucleolar RNA-ended lncRNAs (sno-lncRNAs) expressed from chr15q11-q13 are abundant and accumulate near their sites of synthesis. They sequester Fox2 and might influence Fox2-mediated alternative splicing regulation in *trans* [20].

The mechanism through which these lncRNAs are restricted to their sites of transcription is unknown. Binding to local factors [11–16] or forming RNA–DNA triplex structures [17] may facilitate lncRNAs accumulation in *cis*. Another attractive hypothesis that remains to be tested is that these lncRNAs may remain bound to Pol II [12] and their in *cis* localization is somehow coupled to Pol II transcription.

lncRNAs That Localize in the Nucleus in *trans*

A second group of lncRNAs also localizes to the nucleus, but they often need to be relocated from their sites of synthesis to have an impact on gene regulation, which can be either globally or in a gene-specific manner (Figure 2C). HOTAIR is one of the best-studied examples and can be used to illustrate this group of lncRNAs. HOTAIR is transcribed from the mammalian *HOX C* locus, but it represses transcription in *trans* at the *HOX D* locus by recruiting PRC2 during developmental patterning [21,22]. The expression of HOTAIR also affects the cancer epigenome by genome-wide retargeting of PRC2, leading to altered histone H3 lysine 27 trimethylation (H3K27me3), a chromatin mark that is indicative of transcription repression [23]. In addition to the *HOX D* locus, HOTAIR binds to hundreds of other genomic sites through a GA-rich motif to nucleate broad domains of polycomb occupancy and H3K27me3 [24]. The GA-rich HOTAIR motif may enable formation of an RNA:DNA:DNA triplex that recruits the HOTAIR–chromatin modification complex. BCAR4 and PVT1 are two other lncRNAs that predominately localize in the nucleoplasm. BCAR4 binds to transcription factors SNIP1 and PNUTS via its two distinct regions and alters downstream epigenetic regulation in response to chemokine signals in breast cancer [25]. PVT1 interferes with phosphorylation of MYC at Thr58 in the nucleus, which subsequently increases MYC stability and leads to elevated MYC levels in cancers [26].

Whereas HOTAIR, BCAR4, and PVT1 have a global localization within the nucleus, some other nuclear lncRNAs accumulate within specific nuclear bodies (Figure 2D). These membrane-less subnuclear structures contain specific protein and lncRNA components that define particular nuclear processes. Gomafu (also referred to as MIAT in humans) localizes to distinct nuclear domains in specific neuronal cells and associates with RBPs including Celf3 and SF1 [27], suggesting of role of Gomafu in sequestering proteins. MALAT1 localizes to nuclear speckles, which contain various splicing factors [28], and has been recently shown to regulate mammary cancer pathogenesis [29]. The NEAT1 lncRNAs, expressed from human chr.11q13.1, are abundant lncRNAs that are localized to paraspeckles [30]. There are approximately 5–20 paraspeckles in each nucleus of most mammalian cells and NEAT1 is required for the integrity of paraspeckles [30–33]. NEAT1 interacts with multiple paraspeckle-localized proteins [34,35] and hundreds of genomic sites of active genes in human cells [35]. Although how NEAT1 and paraspeckles are assembled and localized has yet to be defined precisely, it is clear that paraspeckle assembly occurs during the transcription of *NEAT1* [36] and that a change in the transcription of *NEAT1* leads to altered paraspeckle formation [37]. Hence, coupling with Pol II transcription to nucleate nuclear domains represents one way that lncRNAs can nonrandomly accumulate in the nucleus. How other lncRNAs accumulate in *trans* is largely unknown. Since roles of this group of lncRNAs are diverse, mechanisms of their nuclear localization are expected to vary.

lncRNAs That Are Exported to the Cytoplasm

The third group of lncRNAs needs to be exported to the cytoplasm to carry out their regulatory roles (Figure 2E). Increasing lines of evidence show that such lncRNAs can affect gene expression in multiple ways. Cytoplasmic lncRNAs can interfere with protein post-translational modifications, resulting in aberrant signal transduction [38,39]. For example, the lncRNA NKILA interacts with NF-κB and interferes with phosphorylation of IκB. This leads to inhibition of IKK-induced IκB phosphorylation and NF-κB activation and suppression of breast cancer metastasis [38]. Furthermore, cytoplasmic lncRNAs can influence gene regulation by acting as decoys for miRNAs [40] and proteins [41]. It has been shown that linc-MD1 sponges miR-133, which usually targets transcription factors to activate muscle-specific gene expression [40]. A highly conserved and abundant lncRNA NORAD regulates genomic stability by sequestering PUMILIO proteins, which represses the stability and translation of mRNAs to which they bind [41]. Moreover, other lncRNAs can also affect mRNA translation in the cytoplasm [42,43]. STAU1-mediated mRNA decay (SMD) involves the degradation of translationally active mRNAs whose 3′-untranslated regions (3′-UTRs) bind to STAU1, a protein that binds to double-stranded RNA (dsRNA). The requisite dsRNA can be formed by cytoplasmic lncRNAs base pairing with a complementary Alu embedded within the mRNA 3′-UTRs [42]. lincRNA-p21 also has a role in mRNA translation suppression. In the cytoplasm, HuR influences lincRNA-p21 decay; in the absence of HuR, lincRNA-p21 is stable and identifies mRNA targets including JUNB and CTNNB1 mRNAs by base pairing and represses their translation [43].

Finally, recent proteomic analyses revealed that a subset of transcripts currently annotated as lncRNAs may nevertheless encode small proteins (see review [44]) and several annotated lncRNAs can indeed encode functional micropeptides [45–47]. Thus, functional studies of cytoplasmic lncRNAs should also carefully assess whether the lncRNA of interest has any coding potential. In this scenario, one advantage of studying nuclear lncRNAs is that they are excluded from the possibility of being translated, therefore assuring confidence that the experimental conclusions reached truly reflect RNA functions.

Together, these emerging studies reveal that it is critical for lncRNAs to be localized to specific subcellular compartments for their functions. In the next section, the mechanisms that may account for the distinct subcellular fates of different lncRNAs are discussed, focusing on nuclear-retained ones. This area contains many aspects that have not been well addressed in metazoans.

Mechanisms of lncRNA Subcellular Fate Determination

At the chromatin level, lncRNA loci appear to follow the same rules as for protein-coding genes, although there are some differences between chromatin assemblies at divergent promoters for sense mRNA versus antisense lncRNAs in yeast [48]. These similarities led to the systematic discovery of lncRNAs by monitoring the patterns of histone 3 Lys 4 trimethylation (H3K4me3) in the promoter region followed by histone 3 Lys 36 trimethylation (H3K36me3) across the actively transcribed intergenic regions in mammalian cells. The same patterns are also associated with protein-coding genes [3,4,49,50]. These studies indicate that the chromatin status is unlikely to be a major factor in either distinguishing the fate and function of lncRNAs from mRNAs or determining the diverse cellular localizations among lncRNAs. Other than chromatin-mediated regulation, several possible mechanisms may dictate the subcellular localization patterns of lncRNAs.

Nuclear Retention Signals in lncRNAs Are Deposited during Transcription through the Assembly of Long Noncoding Ribonucleoproteins Complexes

An interesting observation is that lncRNAs produced by alternative processing pathways (Box 1 and Figures 1C–G) are likely to be enriched in the nucleus in mammalian cells, suggesting that the processing of lncRNAs may be associated with their distinct subcellular localization and

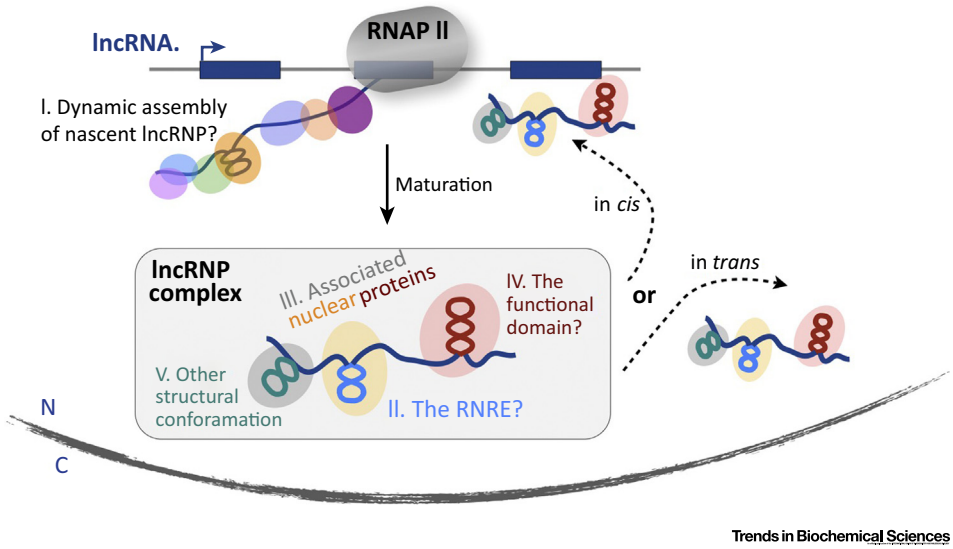

Figure 3. Mechanisms of Long Noncoding RNA (lncRNA) Nuclear Retention. Multiple factors can be involved in lncRNA nuclear retention. The nuclear retention signals of lncRNAs might be deposited during the assembly of nascent long noncoding ribonucleoproteins (lncRNPs) complexes (I). Upon lncRNP maturation, ribonucleic nuclear retention elements (RNREs; II) residing in lncRNAs may associate with nuclear matrix proteins (III, yellow oval) to constrain lncRNAs in the nucleus. In addition, functional domains (IV) of lncRNAs may recruit chromatin-modifying complexes or transcription factors, or other structural conformation of lncRNAs (V) that can subsequently keep these lncRNA molecules from export to the cytoplasm and facilitate their nuclear localization in *cis* or in *trans*. Abbreviation: RNAP, RNA polymerase II.

function, for instance, the lack of poly(A) tail may inhibit lncRNA export. However, whether the transcription and processing of mRNA-like lncRNAs (Figures 1A,B) are linked to their distinct subcellular localization and function is awaiting exploration.

Pol II transcription, nascent pre-mRNA splicing, capping, polyadenylation, mRNA export, and surveillance are seamlessly integrated during mRNA maturation. Throughout this process, pre-mRNAs are decorated with various proteins to form higher-order messenger ribonucleoprotein (mRNP) complexes. Mature mRNPs are then exported to the cytoplasm where translation and mRNA turnover take place (see review [51]).

To distinguish lncRNAs from nuclear-retained mRNAs or cytoplasmic lncRNAs, nuclear retention signals may decorate the nuclear-retained lncRNAs upon transcription. In this model, the sequential assembly of the long noncoding ribonucleoprotein (lncRNP) complexes of future nuclear-retained lncRNAs during their transcription would be substantially distinct from that of mRNPs or that of lncRNPs that needed to be exported to the cytoplasm (Figure 3). Some components of lncRNPs are likely to be different from those of mRNPs. Such dynamic assembly of lncRNPs during transcription and processing may in turn direct their specific subnuclear localization patterns.

Although there is no established evidence yet to support this model in mammalian cells, recent studies in yeast began to reveal differences between the biogenesis of lncRNAs and mRNAs. In the budding yeast *Saccharomyces cerevisiae*, the transcription preinitiation complexes and key transcription factors that regulate lncRNA and mRNA expression appear to have no detectable differences [52]. Rather, later steps in transcription appear to provide a divergence point. By comparing the maturation pathways of mRNAs and different types of lncRNAs, such as stable unannotated transcripts (SUTs) and cryptic unstable transcripts (CUTs), in budding yeast, Tuck and Tollervey [53] observed that the cellular fate of lncRNA and mRNA transcripts is largely

determined during 3′ end formation prior to the acquisition of export competence [53]. The nuclear RNA surveillance machinery rapidly targets CUTs terminated by Nrd1-dependent transcription termination, whereas SUTs contain signals for cleavage and polyadenylation just like mRNAs. Moreover, while most SUTs remain in the nucleus, a subset can be exported to the cytoplasm [53]. However, it remains to be determined whether similar features can be applied to lncRNAs in higher eukaryotic cells. As mRNP transcription, processing, and export are much more complex in mammals, it is likely that the regulation of lncRNA production during transcription and processing in these cells is more complicated than is currently appreciated in yeast. For example, the *CCAT1* locus produces two abundant lncRNAs with overlapping 5′ ends: CCAT1-L (the 4700-nt long isoform) and CCAT1-short isoform (CCAT1-S; the 2600-nt short isoform) that is proposed to be generated by 3′ end processing of the long CCAT1 in colorectal cancers. Strikingly, CCAT1-L accumulates in *cis* in the nucleus while CCAT1-S is mainly cytoplasmic [13], indicating that the prevalence of altered 3′ end formation of Pol II transcripts in mammals [54] may promote different lncRNP assemblies for lncRNA localization.

The coupling of lncRNA transcription and the dynamic assembly of lncRNPs during processing may also be important for both lncRNA subcellular localization and function. For example, the *HAUNT* lncRNA locus appears to regulate nearby *HOX* gene expression through the *HAUNT* genomic DNA, the nuclear-retained HAUNT lncRNA [55], and the transcription of this locus [56]. It will be of interest to examine how the functional outputs of many other lncRNA loci are linked to their own transcription, and whether an alteration of Pol II transcription elongation rate would affect the assembly and subcellular fates of lncRNPs and their potential roles in gene regulation.

Association with Specific Proteins Traps lncRNAs in the Nucleus

Mature lncRNAs may interact with nuclear matrix proteins that would subsequently facilitate nuclear localization of lncRNAs (Figure 3). The matrix protein hnRNP U is required for the accumulation of lncRNA X-inactive specific transcript on the inactive X chromosome [57] and for the in *cis* accumulation of FIRRE [19], which is essential for their functions. Specifically, embryonic stem cells lacking hnRNP U expression fail to form the inactive X chromosome [57] and hnRNP U is essential for FIRRE-mediated *trans*-chromosomal associations at its transcriptional site [19]. In addition, knockdown of hnRNP U even led to a translocation of FIRRE into the cytoplasm [19], indicating that hnRNP U is an important factor for both focal localization and nuclear retention of FIRRE. It remains to be investigated whether hnRNP U, or other unknown nuclear matrix proteins, can serve as common factors that are required for in *cis* accumulation or nuclear retention of lncRNAs. Such proteins would have the potential to be essential for both lncRNA subcellular localization and function. Indeed, lncRNAs might be retained in the nucleus through the formation of a mature and functional lncRNP complex including chromatin-modifying complexes or transcription factors. For instance, the interaction of nucleus-localized chromatin-modifying complexes with HOTAIR [21,22] and transcription factors with BCAR4 [25] may in turn keep these lncRNAs from nuclear export.

The higher-order organization of chromosomes may also promote nuclear retention of lncRNAs by specifying their target sites distantly. For instance, the lncRNA HOTTIP is expressed from the 5′ end of the *HoxA* locus; it is brought to the *HoxA* distal genes by chromosomal looping to activate its target genes at a distance through binding the WDR5/MLL complex [58]. One speculation as to how lncRNAs are relocated to distinct sites in the nucleus is that the dynamic remodeling of lncRNPs during the lifetime of an lncRNA together with the 3D chromosomal configuration may facilitate this nuclear transportation. Future studies using genome-editing technologies to knock-in RNA aptamers at the endogenous lncRNA locus of interest followed by live cell imaging may provide substantially novel insights into lncRNA or lncRNP transport within the nucleus.

Nuclear Retention Signals Embedded within lncRNAs Constrain lncRNAs in the Nucleus

The lncRNA itself may possess essential primary sequences, or more likely secondary or tertiary structures, that are required for nuclear localization (Figure 3). A nuclear retention element containing a U-rich internal loop that interacts with the 3′ poly(A) tail to stabilize nuclear-accumulated RNAs has been identified near the 3′ end of the nuclear-restricted polyadenylated PAN lncRNA from Kaposi's sarcoma-associated herpesvirus [59,60]. A similar structure is in the nuclear-retained MALAT1 and the long isoform of NEAT1 [60,61], but it is undetermined whether such an element acts for nuclear retention of these lncRNAs. Furthermore, a pentamer RNA motif residing in the lncRNA BORG is important for its proper nuclear localization [62]. Similarly, FIRRE has a unique highly structured 156-bp repeat named the repeating RNA domain (RRD), and FIRRE lacking the RRD fails to accumulate on the chromosome [19]. Intron-containing circular RNAs [63,64] and intron-derived sno-lncRNAs [20] are retained in the nucleus, indicating that intronic sequences may serve as ribonucleic nuclear retention elements (RNREs) that lead to lncRNA nuclear localization. Finally, some divergently transcribed lncRNAs can form DNA–RNA triplexes to anchor the lncRNAs to the vicinity of their gene loci [17].

These studies together reveal that RNA sequences can act as RNREs that mediate nuclear localization. Remarkably, such RNREs are diverse in terms of their length and sequence conservation and how RNREs promote nuclear retention of lncRNAs remains elusive. Due to the low sequence conservation of many lncRNAs, the structural domains of only a few lncRNAs have been solved. One example is the lncRNA Gas 5, which inhibits the transcriptional activity of steroid receptors (SRs) through direct competition for DNA binding. A putative 33 base stem–loop region of Gas5 has been shown to 'mimic' genomic SR-binding sites, therefore forming a compact lncRNA (Gas5)–protein (SR) interaction domain [65]. A more recent study using chemical probing and phylogenetic analysis revealed that the structural organization of HOTAIR was comparable to well-folded RNAs, like Group II introns and Gas5. The secondary structure of HOTAIR is composed of four independently folding modules corresponding to its predicted protein-binding domains [66]. It is unclear whether such folding modules can be served as RNREs for HOTAIR.

Finally, multiple factors including RNREs, higher-order chromosome organization, nuclear protein factors, and the coupling of RNP assembly with lncRNA transcription may often function together to achieve the unique subcellular localization of specific lncRNAs. The application of different approaches (Box 2) may be capable of revealing the mechanism that links lncRNA localization and function.

lncRNAs in the Cytoplasm

Since cytoplasmic lncRNAs and mRNAs appear similar in many aspects, an obvious hypothesis is that the export of such lncRNAs is regulated by mechanisms similar to those that regulate mRNA export, although the associated proteins for lncRNAs and mRNAs could be distinct. For example, the association of exon junction complexes (EJCs) with lncRNAs containing splice junction sites could be substantially different from that of mRNAs, as EJCs are known to be deposited on newly spliced mRNAs in the nucleus and coordinate several downstream steps of mRNA biogenesis, including nuclear export and translation (see reviews [51,67]). Interestingly, if EJCs are deposited on spliced cytoplasmic lncRNAs, then it is likely that cytoplasmic lncRNPs are not remodeled like mRNPs, which are changed dramatically by the act of translation [67]. For example, cytoplasmic lncRNAs do not go through the pioneer round of translation and they may not be remodeled the same way that mRNPs are. If true, then cytoplasmic lncRNPs may have unique and identifying RNP structures that would allow some lncRNAs escape from polyribosome trapping and mRNA turnover pathways.

Box 2. Emerging Approaches to Study Long Noncoding RNA (lncRNA) Localization and Function.

Visualization of lncRNAs in fixed cells is readily accomplished through regular or single-molecule RNA fluorescence *in situ* hybridization [10]. A new method that examines endogenous RNA in living cells uses the nuclear-localized RNA-targeting CRISPR-Cas9 system [82]. Application of native elongating transcript sequencing, which aims to study the complex events underlying transcription in mammalian cells [83], may uncover transcriptional differences between lncRNA genes that have different subcellular fates. RNA antisense purification methods enable the analysis of lncRNA-associated proteins, RNAs, and DNAs with much higher resolution. These methods have provided new mechanisms of X chromosome inactivation by identifying XIST-interacting genome architectures and proteins [84–86] and have also defined new genomic targets of NEAT1 [35] and MALAT1 [35,87]. Detailed studies are still needed to dissect which proteins are necessary for the localization or function of each lncRNA. Deciphering the structural domains of each lncRNA will be crucial to solving the specificity of long noncoding ribonucleoprotein (lncRNP) interactions. A number of reagents have been used to probe RNA structures both at the individual level in a conventional manner and at the genome-wide level through combination with deep sequencing (see review [7]). However, such studies are still far from sufficient to understand the precise molecular basis of how individual structural motifs within lncRNPs act. New methodologies and algorithms may enable the solution of structures of defined RNA structural domains with their interacting proteins. For example, small-angle X-ray scattering (SAXS) is increasingly used to characterize the structure and interactions of macromolecules in solution. SAXS has revealed that the HIV Rev response element, a 232-nt-long RNA motif that the virus uses to recognize its own mRNA among more abundant host cellular RNAs, has an unusual topological structure [88]. Electron cryomicroscopy has also recently been shown to have a powerful capability to reveal the structures of large RNA–protein complexes captured in various states, which has provided insights into yeast spliceosome assembly and catalysis [89,90]. The application of similar approaches to study lncRNP complexes will dissect the essential structures of lncRNAs and their interacting protein complexes in detail.

Outstanding Questions

How is the specific localization of long noncoding RNAs (lncRNAs) achieved and regulated?

Does the dynamic assembly of long noncoding ribonucleoproteins determine their subcellular localization?

How do the RNA structural motifs and effective protein partners work together to define the cellular fates of lncRNAs?

What is the life cycle of lncRNAs in the cytoplasm?

What approaches could be applied to resolve lncRNA structural domains and lncRNA–protein complexes in detail?

Concluding Remarks

The lncRNAs are a large and diverse class of transcripts that affect gene regulation through a variety of mechanisms. Depending on their genomic origin, subcellular localization, or functional pathways, lncRNAs can be classified into different groups. Like proteins, lncRNAs must localize to specific subcellular compartments to execute their functions. However, how the specific localization of lncRNAs is achieved and regulated and what rules lncRNAs follow to make them so remarkably different from mRNAs remain largely unknown. In addition to the unexamined dynamic assembly of lncRNPs during transcription and processing, other factors including higher-order nuclear organization, RNA structural motifs, and protein partners are also proposed to play important roles in defining lncRNA subcellular localization. It will be of great interest to identify common RNA structural domains and protein factors that regulate the nuclear retention or accumulation of lncRNAs. In addition, the nucleocytoplasmic export of cytoplasmic lncRNAs and their life cycle in the cytoplasm also require a thorough investigation. Nevertheless, understanding these features of lncRNAs will greatly expand our current knowledge of lncRNA biology and shed new lights into the study of their cellular roles in depth (see Outstanding Questions).

Acknowledgments

I thank Li Yang, Gordon Carmichael, and members of the Chen lab for critical reading the manuscript and apologize to colleagues whose works are not discussed due to space limitation. This work was supported by grants 91440202 and 31322018 from NSFC and 2016YFA0100701 from MOST.

References

1. Quinn, J.J. and Chang, H.Y. (2015) Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62
2. Rinn, J.L. and Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166
3. Goff, L.A. and Rinn, J.L. (2015) Linking RNA biology to lncRNAs. *Genome Res.* 25, 1456–1465
4. Guttman, M. and Rinn, J.L. (2012) Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339–346
5. Cabili, M.N. *et al.* (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Gene Dev.* 25, 1915–1927
6. Katayama, S. *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science* 309, 1564–1566
7. Chu, C. *et al.* (2015) Technologies to probe functions and mechanisms of long noncoding RNAs. *Nat. Struct. Mol. Biol.* 22, 29–35
8. Batista, P.J. and Chang, H.Y. (2013) Long noncoding RNAs: cellular address codes in development and disease. *Cell* 152, 1298–1307
9. Chen, L.L. and Carmichael, G.G. (2010) Decoding the function of nuclear long non-coding RNAs. *Curr. Opin. Cell Biol.* 22, 357–364
10. Cabili, M.N. *et al.* (2015) Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* 16, 20
11. Wang, X. *et al.* (2008) Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* 454, 126–130

12. Sigova, A.A. *et al.* (2015) Transcription factor trapping by RNA in gene regulatory elements. *Science* 350, 978–981
13. Xiang, J.F. *et al.* (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* 24, 513–531
14. Nagano, T. *et al.* (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 322, 1717–1720
15. Terranova, R. *et al.* (2008) Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev. Cell* 15, 668–679
16. Pandey, R.R. *et al.* (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* 32, 232–246
17. Postepska-Igielska, A. *et al.* (2015) LncRNA Khps1 regulates expression of the proto-oncogene SPHK1 via triplex-mediated changes in chromatin structure. *Mol. Cell* 60, 626–636
18. Latos, P.A. *et al.* (2012) Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338, 1469–1472
19. Hacisuleyman, E. *et al.* (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206
20. Yin, Q.F. *et al.* (2012) Long noncoding RNAs with snoRNA ends. *Mol. Cell* 48, 219–230
21. Rinn, J.L. *et al.* (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311–1323
22. Tsai, M.C. *et al.* (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693
23. Gupta, R.A. *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076
24. Chu, C. *et al.* (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* 44, 667–678
25. Xing, Z. *et al.* (2014) lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. *Cell* 159, 1110–1125
26. Tseng, Y.Y. *et al.* (2014) PVT1 dependence in cancer with MYC copy-number increase. *Nature* 512, 82–86
27. Ishizuka, A. *et al.* (2014) Formation of nuclear bodies by the lncRNA Gomafu-associating proteins Celf3 and SF1. *Genes Cells* 19, 704–721
28. Tripathi, V. *et al.* (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938
29. Arun, G. *et al.* (2016) Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Gene Dev.* 30, 34–51
30. Clemson, C.M. *et al.* (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* 33, 717–726
31. Sunwoo, H. *et al.* (2009) MEN epsilon/beta nuclear-retained noncoding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* 19, 347–359
32. Chen, L.L. and Carmichael, G.G. (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* 35, 467–478
33. Sasaki, Y.T. *et al.* (2009) MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2525–2530
34. Naganuma, T. *et al.* (2012) Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* 31, 4020–4034
35. West, J.A. *et al.* (2014) The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* 55, 791–802
36. Mao, Y.S. *et al.* (2011) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* 13, 95–101
37. Hu, S.B. *et al.* (2015) Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus. *Genes Dev.* 29, 630–645
38. Liu, B. *et al.* (2015) A cytoplasmic NF-κB interacting long noncoding RNA blocks IκB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 27, 370–381
39. Lin, A. *et al.* (2016) The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer. *Nat. Cell Biol.* 18, 213–224
40. Cesana, M. *et al.* (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358–369
41. Lee, S. *et al.* (2016) Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* 164, 69–80
42. Gong, C. and Maquat, L.E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* 470, 284–288
43. Yoon, J.H. *et al.* (2012) LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* 47, 648–655
44. Cohen, S.M. (2014) Everything old is new again: (linc)RNAs make proteins! *EMBO J.* 33, 937–938
45. Anderson, D.M. *et al.* (2015) A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 160, 595–606
46. Lauressergues, D. *et al.* (2015) Primary transcripts of microRNAs encode regulatory peptides. *Nature* 520, 90–93
47. Nelson, B.R. *et al.* (2016) A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* 351, 271–275
48. Marquardt, S. *et al.* (2014) A chromatin-based mechanism for limiting divergent noncoding transcription. *Cell* 157, 1712–1723
49. Khalil, A.M. *et al.* (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11667–11672
50. Ulitsky, I. *et al.* (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537–1550
51. Chuang, T.W. *et al.* (2015) Function and pathological implications of exon junction complex factor Y14. *Biomolecules* 5, 343–355
52. Rhee, H.S. and Pugh, B.F. (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* 483, 295–301
53. Tuck, A.C. and Tollervey, D. (2013) A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* 154, 996–1009
54. Tian, B. and Manley, J.L. (2013) Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem. Sci.* 38, 312–320
55. Maamar, H. *et al.* (2013) Linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in *cis*. *Genes Dev.* 27, 1260–1271
56. Yin, Y. *et al.* (2015) Opposing roles for the lncRNA Haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Cell Stem Cell* 16, 504–516
57. Hasegawa, Y. *et al.* (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* 19, 469–476
58. Wang, K.C. *et al.* (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124
59. Mitton-Fry, R.M. *et al.* (2010) Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 330, 1244–1247
60. Brown, J.A. *et al.* (2012) Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19202–19207
61. Wilusz, J.E. *et al.* (2012) A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* 26, 2392–2407
62. Zhang, B. *et al.* (2014) A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. *Mol. Cell. Biol.* 34, 2318–2329

63. Zhang, Y. *et al.* (2013) Circular intronic long noncoding RNAs. *Mol. Cell* 51, 792–806
64. Li, Z. *et al.* (2015) Exon-intron circular RNAs regulate transcription in the nucleus. *Nat. Struct. Mol. Biol.* 22, 256–264
65. Hudson, W.H. *et al.* (2014) Conserved sequence-specific lincRNA-steroid receptor interactions drive transcriptional repression and direct cell fate. *Nat. Commun.* 5, 5395
66. Somarowthu, S. *et al.* (2015) HOTAIR forms an intricate and modular secondary structure. *Mol. Cell* 58, 353–361
67. Bono, F. and Gehring, N.H. (2011) Assembly, disassembly and recycling: the dynamics of exon junction complexes. *RNA Biol.* 8, 24–29
68. Hutchinson, J.N. *et al.* (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* 8, 39
69. Dhir, A. *et al.* (2015) Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs. *Nat. Struct. Mol. Biol.* 22, 319–327
70. Kim, T.K. *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187
71. De Santa, F. *et al.* (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* 8, e1000384
72. Lam, M.T. *et al.* (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* 498, 511–515
73. Li, W. *et al.* (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520
74. Melo, C.A. *et al.* (2013) eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell* 49, 524–535
75. Lai, F. *et al.* (2015) Integrator mediates the biogenesis of enhancer RNAs. *Nature* 525, 399–403
76. Preker, P. *et al.* (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 322, 1851–1854
77. Preker, P. *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
78. Ntini, E. *et al.* (2013) Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat. Struct. Mol. Biol.* 20, 923–928
79. Lubas, M. *et al.* (2015) The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.* 10, 178–192
80. Zhang, X.O. *et al.* (2014) Species-specific alternative splicing leads to unique expression of sno-lncRNAs. *BMC Genomics* 15, 287
81. Chen, L.L. (2016) The biogenesis and emerging roles of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 17, 205–211
82. Nelles, D.A. *et al.* (2016) Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell* 165, 488–496
83. Nojima, T. *et al.* (2015) Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 161, 526–540
84. Engreitz, J.M. *et al.* (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341, 1237973
85. McHugh, C.A. *et al.* (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236
86. Chu, C. *et al.* (2015) Systematic discovery of Xist RNA binding proteins. *Cell* 161, 404–416
87. Engreitz, J.M. *et al.* (2014) RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell* 159, 188–199
88. Fang, X. *et al.* (2013) An unusual topological structure of the HIV-1 Rev response element. *Cell* 155, 594–605
89. Yan, C. *et al.* (2015) Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* 349, 1182–1191
90. Wan, R. *et al.* (2016) The 3.8 Å structure of the U4/U6.U5 tri-snRNP: insights into spliceosome assembly and catalysis. *Science* 351, 466–475