Linking RNA Processing and Function

Run-Wen Yao,[1] Chu-Xiao Liu,[1] and Ling-Ling Chen[1,2]

[1]*State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China*

[2]*School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China*

Correspondence: linglingchen@sibcb.ac.cn

RNA processing is critical for eukaryotic mRNA maturation and function. It appears there is no exception for other types of RNAs. Long noncoding RNAs (lncRNAs) represent a subclass of noncoding RNAs, have sizes of >200 nucleotides (nt), and participate in various aspects of gene regulation. Although many lncRNAs are capped, polyadenylated, and spliced just like mRNAs, others are derived from primary transcripts of RNA polymerase II and stabilized by forming circular structures or by ending with small nucleolar RNA–protein complexes. Here we summarize the recent progress in linking the processing and function of these unconventionally processed lncRNAs; we also discuss how directional RNA movement is achieved using the radial flux movement of nascent precursor ribosomal RNA (pre-rRNA) in the human nucleolus as an example.

Surprisingly, the human genome is pervasively transcribed (>80%), and >98% of this transcriptional output represents non-protein-coding RNAs (ncRNAs) (Derrien et al. 2012; Mudge et al. 2013). Long ncRNAs (lncRNAs), which are longer than 200 nucleotides (nt) and lack protein-coding potential, have emerged as a major class of eukaryotic regulatory transcripts involved in multiple layers of gene expression. Statistics from Human GENCODE Release version 28 suggest that the human genome contains more than 16,000 lncRNA genes, but other estimates for the number of lncRNA genes exceed 100,000 in humans (Zhao et al. 2016).

Over the past decades, the study of lncRNA biogenesis and regulation has greatly improved our understanding of the overall regulatory RNA diversity and function, ranging from the classical mRNA-like lncRNAs (e.g., *Xist* [Engreitz et al. 2013; Simon et al. 2013; Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015; Chen et al. 2016; Creamer and Lawrence 2017], *CCAT1-L* [Xiang et al. 2014], *NORAD* [Lee et al. 2016; Tichon et al. 2016, 2018; Munschauer et al. 2018]) to other types of unconventionally formatted linear (e.g., *NEAT1* [Clemson et al. 2009; Sunwoo et al. 2009; Souquere et al. 2010; Adriaens et al. 2016; Wang et al. 2018], *MALAT1* [Wilusz et al. 2008; Tripathi et al. 2010; Nakagawa et al. 2012; Zhang et al. 2012; Arun et al. 2016; Malakar et al. 2017], *sno-lncRNA*s [Yin et al. 2012; Xing et al. 2017], *SPA*s [Wu et al. 2016; Lykke-Andersen et al. 2018]) and circular RNAs (Salzman et al. 2012; Hansen et al. 2013; Jeck et al. 2013; Memczak et al. 2013; Zhang et al. 2013). Compared to microRNAs, which are ~22-nt RNAs that mainly direct post-transcriptional repression of mRNA targets in eukaryotes (Bartel 2018), lncRNAs exhibit a surprisingly wide range of sizes, shapes, and functions (Wu et al. 2017; Kopp and Mendell 2018; Uszczynska-Ratajczak et al. 2018; Carlevaro-Fita and Johnson 2019; Yao et al. 2019a). It is now well-known that lncRNAs participate in the regulation of genetic flow of protein expression from chromatin organization to transcription regulation in the nucleus to modulation of mRNA stability, translation, and post-translation in the cytoplasm. These diverse functional potentials depend on the processing, subcellular localization, and formation of structural modules of individual lncRNAs to partner with associated proteins, which may undergo rapid changes depending on local or cellular environments.

Most annotated lncRNAs transcribed from genomic intergenic regions (lincRNAs) by RNA polymerase II (Pol II) are capped, polyadenylated, and spliced just like mRNAs (Cabili et al. 2011; Derrien et al. 2012; Djebali et al. 2012; Quinn and Chang 2016). Although exhibiting tissue- or cell type–specific expression (Cabili et al. 2011; Goff and Rinn 2015; Ulitsky 2016), they differ from mRNAs by being less evolutionarily conserved and less abundant and containing fewer exons. lincRNAs are generally more nuclear localized than their mRNA counterparts (Derrien et al. 2012; Cabili et al. 2015), in part because some lincRNAs are transcribed by deregulated Pol II, weakly spliced (Mele et al. 2017), inefficiently polyadenylated, and degraded by the nuclear exosome on chromatin (Schlackow et al. 2017). Thus, in principle, functional lincRNAs must escape this targeted nuclear surveillance process to accumulate to high levels in specific cell types.

In addition to lincRNAs, recent studies have uncovered other types of lncRNAs that are processed from primary Pol II transcripts and are stabilized by distinct mechanisms. In this review, we briefly summarize our current understanding of the biogenesis of these unconventionally processed lncRNAs and emphasize the importance of

linking their processing to functions in innate immunity and in the regulation of structure and function of human nuclear subdomains. To be functional, these different types of lncRNAs need to be translocated to different subcellular compartmentations. Although mechanisms of RNA trafficking remain incompletely understood, we discuss some insights into this question using the radial flux nature of nascent precursor ribosomal RNA (pre-rRNA) trafficking in the nucleolus as a model.

MALAT1, NEAT1, AND NONPOLYADENYLATED TRANSCRIPTOMES

Earlier studies from the Spector laboratory reported alternative 3′-end processing of MALAT1 (Wilusz et al. 2008) and NEAT1_long (the long isoform of NEAT1 [Sunwoo et al. 2009]), which are two abundant nuclear-enriched lncRNAs that are localized to nuclear speckles and paraspeckles, respectively (Hutchinson et al. 2007; Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009). These RNAs are processed at their 3′ ends by recognition and cleavage of tRNA-like structures by RNase P (which is best known to process the 5′ ends of tRNAs) (Wilusz et al. 2008; Sunwoo et al. 2009). RNase P cleavage leads to the formation of 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) (Wilusz et al. 2012; Brown et al. 2012). A similar triple-helical structure, but not formed by RNase P processing and called a nuclear retention element (NRE), has also been found at the 3′ end of the PAN lncRNA, which is expressed from Kaposi's sarcoma–associated herpesvirus (KSHV), and in RNAs from other viruses (Mitton-Fry et al. 2010; Tycowski et al. 2012). In addition to RNase P–mediated 3′ processing of lncRNAs, a group of lncRNA transcripts containing miRNAs (lnc-pri-miRNAs) use Microprocessor cleavage to terminate transcription, resulting in unstable lncRNAs without 3′-end poly(A) tails (Dhir et al. 2015).

Inspired by alternative 3′-end processing of MALAT1 and NEAT1, we began to explore the nonpolyadenylated (poly(A)) transcriptomes in human cells (Yang et al. 2011). In this method, total RNAs collected from human cells were incubated with oligo(dT) beads to select poly(A) RNAs. The unbound, flow-through RNAs from the oligo(dT) beads were collected and subjected to multiple rounds of rRNA depletion. Both poly(A) RNAs and poly(A)-/ribo-RNAs were then subjected to RNA-seq library preparation and sequencing (Yang et al. 2011). In addition to the replication-dependent histone mRNAs, this non-poly(A) RNA-seq unexpectedly identified hundreds of abundant RNA signals that did not align fully to annotated genes but derived from either introns and exons, which were termed as "excised introns" and "excised exons" (Yang et al. 2011). Later, these excised "exons" were identified as circular RNAs produced from backsplicing of exons of pre-mRNAs (Salzman et al. 2012; Zhang et al. 2014a); "excised introns" were characterized as circular

intronic RNAs (ciRNAs) (Zhang et al. 2013) and snoRNA-ended lncRNAs (sno-lncRNAs) (Yin et al. 2012). Further, SPA (5′ snoRNA capped and 3′ polyadenylated) lncRNAs were uncovered using fibrillarin (a key protein component of Box C/D snoRNP complex)—RNA immunoprecipitation (RIP) and RNA-seq (Wu et al. 2016).

IDENTIFICATION OF DIFFERENT TYPES OF CIRCULAR RNAs FROM NON-POLY(A) TRANSCRIPTOMES

Covalently closed circular RNAs were first observed by electron microscopy in plants and eukaryotic cells in the 1970s (Sanger et al. 1976; Hsu and Cocaprados 1979). This was followed by observations of RNA circles with scrambled exons, which were thought to be "by-products" of aberrant splicing with little functional potential (Nigro et al. 1991; Cocquerelle et al. 1992; Capel et al. 1993).

Genome-wide studies of rRNA-depleted and non-poly (A) transcriptomes, as well as rRNA-depleted and RNase R–enriched transcriptomes, together with specific bioinformatics algorithms for circular RNA detection (Kristensen et al. 2019), have revealed widespread expression of circular RNAs in metazoans (Yang et al. 2011; Gardner et al. 2012; Salzman et al. 2012, 2013; Jeck et al. 2013; Memczak et al. 2013; Zhang et al. 2013, 2014a; Guo et al. 2014; Westholm et al. 2014; Ivanov et al. 2015) and in plants (Wang et al. 2014; Lu et al. 2015). These include two types of circular RNAs derived from the primary Pol II transcripts.

The first type of ciRNAs derive from lariat introns in mammalian cells, and their formation depends on a consensus sequence containing a 7-nt GU-rich motif near the 5′ splice site and an 11-nt C-rich motif at the branchpoint site that inhibits debranching. The resulting RNA circles are covalently ligated through a 2′, 5′-phosphodiester bond at the joining site and lack the linear part stretching from the 3′ end of the intron to the branchpoint (Fig. 1A). Hundreds of ciRNAs were identified in human cells including the human embryonic carcinoma cell line PA1 and the human embryonic stem cell (hESC) H9 line. Some abundantly expressed ciRNAs including ci-ankrd52 and ci-sirt7 were found to primarily accumulate in the nucleus and interact with the elongating Pol II complex. Depleting these ciRNAs led to decreased transcription of the parental ANKRD52 or SIRT7 genes. These results suggest a cis-regulatory role of intronic noncoding sequences on their parent coding genes (Fig. 1A; Zhang et al. 2013). Stable intronic RNAs derived from lariats were also found in oocytes of Xenopus tropicalis (Gardner et al. 2012; Talhouarne and Gall 2014) and Drosophila melanogaster (Jia Ng et al. 2018).

Of note, different from higher eukaryotic cells, it has recently seen that excised introns in yeast can be stabilized as linear forms that can regulate cell growth (Morgan et al. 2019; Parenteau et al. 2019). Thirty-four excised linear introns were found in Saccharomyces cerevisiae that remain associated with components of the spliceosome. These differ from classical spliceosomal introns by con-
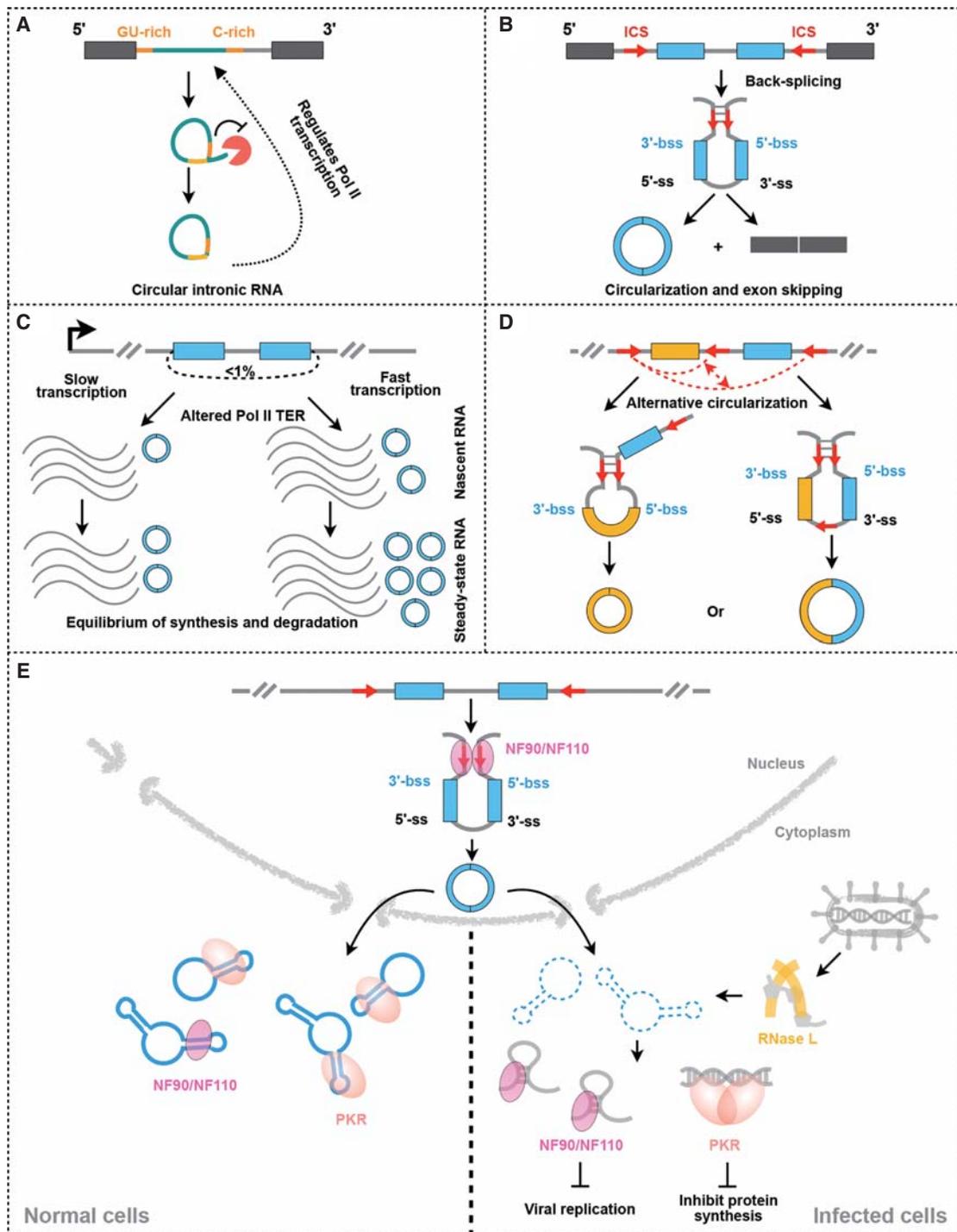
Figure 1. Production, structure, and degradation of circular RNAs. (*A*) A model for circular intronic RNA (ciRNA) processing. ciRNAs are derived from lariat introns and are covalently ligated via a 2′,5′-phosphodiester bond. Their formation depends on a consensus sequence near the 5′ splice site and containing a 7-nt GU-rich motif near the 5′ splice site and an 11-nt C-rich motif at the branchpoint site. (*B*) A model for circular RNA (circRNA) processing. circRNAs are produced by spliceosome-catalyzed backsplicing of exon(s). Generally, flanking intronic complementary sequences (ICSs; most are *Alu* elements) promote exon circularization, whose efficiency can be regulated by competition between ICSs across introns. (*C*) Nascent circRNA production correlates with Pol II elongation rate. In cells, the efficiency of backsplicing from pre-mRNA is relatively low, but because of the stability, circRNAs can accumulate to achieve high abundance at the steady state levels. (*D*) A model of alternative circularization. Alternative formation of inverted repeated ICS pairs and the competition between the pairs can lead to multiple circRNAs produced from a single gene locus. (The red dashed arrows indicate potential competition between different ICS pairs.) (*E*) Linking circRNA processing and structure to innate immunity regulation. Under normal conditions (*left*), NF90/NF110 promotes circRNA production by stabilizing ICSs flanking intronic RNA pairs to juxtapose the circRNA-forming exon(s). Many examined circRNAs form intramolecular short imperfect double-stranded RNAs (dsRNAs) that act as inhibitors of innate immune dsRNA receptor PKR and NF90/NF110. Upon viral infection (*right*), circRNAs are globally reduced at the production level because of export of NF90/NF110, as well as at the steady state level because of the rapid turnover upon RNase L activation. This global reduction of circRNAs may free PKR and NF90/NF110 to be engaged in antiviral immune responses. Misregulation of this process is found in patients with autoimmune disease including systemic lupus erythematosus (SLE).

taining a short distance between their lariat branchpoint and the 3′ splice site (Morgan et al. 2019). Such linear but stable introns were shown to promote resistance to starvation by enhancing the repression of ribosomal protein genes that are downstream from the nutrient-sensing TORC1 and protein kinase A (PKA) pathways (Morgan et al. 2019).

A second type of circular RNAs (circRNAs) are produced from backsplicing of exons of pre-mRNAs in thousands of gene loci in eukaryotes. For these, exons from pre-mRNAs have their ends covalently joined via 3′, 5′-phosphodiester bonds. Although the expression level of most circRNAs is low, some circRNAs have been reported to accumulate to levels as high as or even higher than that of their linear cognate mRNAs (Salzman et al. 2013; Rybak-Wolf et al. 2015; You et al. 2015; Zhang et al. 2016b). Their biogenesis requires spliceosomal machinery and can be modulated by both *cis*- and *trans*-factors. At the molecular level, some abundant circRNAs were shown to modulate gene expression by titrating miRNAs (Hansen et al. 2013; Memczak et al. 2013; Piwecka et al. 2017) or proteins (Chen et al. 2017; Li et al. 2017; Xia et al. 2018; Liu et al. 2019), regulating transcription (Li et al. 2015; Conn et al. 2017), interfering with splicing (Ashwal-Fluss et al. 2014; Zhang et al. 2014a, 2016a), or even acting as templates for translation (Legnini et al. 2017; Pamudurti et al. 2017; Yang et al. 2017). However, because of their circular conformation and sequence overlap with linear mRNA counterparts, challenges exist at multiple levels, from annotations to functional studies, to understand the expression and functions of circRNAs (Li et al. 2018). In the past several years, we have focused on understanding the regulation of their life cycles, which led us to uncover important functions in innate immune responses (see below).

CHARACTERIZATION OF circRNA BIOGENESIS

During exon backsplicing, a downstream splice-donor site is covalently ligated to an upstream splice-acceptor site (Salzman et al. 2012; Jeck et al. 2013; Memczak et al. 2013; Salzman et al. 2013). How does the spliceosome overcome the sterically unfavorable reaction between backspliced exons? The detailed mechanism of circRNA biogenesis had remained to be explored, despite a noted association with *Alu* elements or other complementary sequences in flanking introns of circle-forming exons (Capel et al. 1993; Jeck et al. 2013). By developing a computational algorithm (CIRCexplorer) (Zhang et al. 2014a), we identified thousands of circRNAs in non-poly (A) data sets derived from H9 and HeLa cells [described as "excised exons" (Yang et al. 2011)]. Using circular RNA recapitulation assays that contain flanking intronic complementary sequences in circle-forming vectors, we have shown that exon circularization is in general promoted by flanking intronic complementary sequences (ICSs), most being *Alu* elements in human (Fig. 1B). The efficiency of exon circularization is regulated by competition between

RNA pairing across flanking introns or within individual introns (Zhang et al. 2014a). Given that the great majority of circRNAs are derived from the middle exons of genes (Zhang et al. 2014a) and that backsplicing requires the canonical spliceosomal machinery (Ashwal-Fluss et al. 2014; Starke et al. 2015; Wang and Wang 2015; Liang et al. 2017), it has been proposed that backsplicing events compete with linear RNA splicing (Fig. 1B; Ashwal-Fluss et al. 2014; Zhang et al. 2014a). It would therefore be important to address the kinetics of backsplicing in cells.

To understand how circRNA biogenesis is linked to transcription and canonical splicing, we studied circRNA processing using metabolic tagging of nascent RNAs via 4-thiouridine (4sU), followed by purification of labeled nascent RNAs and RNA-seq (4sUDRB-seq). We found that backsplicing from pre-mRNA is inefficient in cells (<1% of canonical splicing) and that many backsplicing events occur post-transcriptionally (Fig. 1C). However, circRNAs are stable and some can accumulate to high steady state levels (Fig. 1C). For example, in neurons that have slow division rates, we observed remarkably increased number and abundance of circRNAs upon human embryonic stem (ES) cells differentiation to forebrain neurons (Zhang et al. 2016b).

Despite slow processing of circRNAs in cells, the potential alternative formation of inverted repeated ICS pairs and the competition between these pairs lead to the production of multiple circular RNAs from a single gene, a phenomenon that we termed as "alternative circularization" (Fig. 1D). Alternative circularization widely occurs; for example, >50% of highly expressed circRNA (mapped backsplice junction reads ≥ 0.1 RPM [reads per million mapped reads]) gene loci in multiple examined cell lines can produce more than one circRNA (Zhang et al. 2014a; Zhang et al. 2016a). We have annotated patterns of alternative circularization, including different types of alternative backsplicing and alternative splicing events, in circRNAs from a range of commonly used human cell lines (Zhang et al. 2016a). Compared to linear cognate RNAs, circRNAs exhibit distinct patterns of alternative backsplicing and alternative splicing. Quantification of RNA pairing capacity of orientation-opposite ICSs across circRNA-flanking introns using a complementary sequence index revealed that among all types of complementary sequences, short interspersed nuclear elements (SINEs)—especially *Alu* elements in human—contribute the most for circRNA formation, and that their diverse distribution across species results in the increased complexity of circRNA expression during species evolution (Dong et al. 2017). These findings together reveal the complexity of post-transcriptional regulation in mammalian transcriptomes.

PRODUCTION, STRUCTURE, AND DEGRADATION OF circRNAs REGULATE INNATE IMMUNE RESPONSES

Because of overlapping sequences, it has been challenging to study individual circRNA function due to in-

adequate methods in distinguishing exons in circRNAs from those in linear cognate mRNAs (Li et al. 2018). We speculated that understanding the life cycle difference between circular and linear RNAs might provide some insights into their functions.

Although having the same *cis*-elements, expression levels of circRNAs from the same loci exhibit cell type– and tissue-specific patterns, indicating the participation of protein factors in circRNA biogenesis. To identify such factors, we applied a genome-wide siRNA screening that targets all human unique genes with an in-house-developed efficient dual-color circular/linear RNA expression reporter. In this vector, backsplicing produces a translatable circular *mCherry* and canonical splicing produces a translatable linear *egfp*. This screening identified 103 proteins that have an impact on mCherry expression but not EGFP (Li et al. 2017). In addition to protein candidates that have known roles in splicing regulation, we identified multiple factors related to host immune responses involved in circRNA production (Li et al. 2017). One such factor is the human interleukin enhancer binding factor 3 (ILF3), whose gene produces multiple mRNA isoforms by alternative splicing to express factors including nuclear factor 90 (NF90) and nuclear factor 110 (NF110). iCLIP and mutagenesis experiments confirmed that NF90/NF110 directly bind to inverted-repeated *Alu* elements juxtaposing circRNA-forming exon(s) to promote circRNA production by stabilizing intronic RNA pairs in the nucleus (Fig. 1E). Interestingly, mature circRNAs as a group were found to associate with NF90/NF110 in the cytoplasm (Li et al. 2017). It is known that upon viral infection, NF90/NF110 are rapidly exported to the cytoplasm, where they participate in innate immunity (Harashima et al. 2010). After infection, we observed a global reduction of nascent circRNA production, which was consistent with a deassociation of NF90/NF110 from circRNA–protein complexes for antiviral activity. Consistently, overexpression of endogenous circRNAs facilitated viral infection of human cells (Li et al. 2017). These findings indicated that circRNA biogenesis is unfavorable for innate immune responses, thus linking endogenous circRNA production to innate immunity regulation (Fig. 1E; Li et al. 2017).

Remarkably, not only is nascent circRNA production limited globally, but the steady state level of circRNAs is also globally and rapidly (with a turnover of ~1 h) reduced upon poly(I:C) treatment to mimic pathogenic dsRNAs or the encephalomyocarditis virus (EMCV) infection (Liu et al. 2019). This global reduction of steady state circRNAs is catalyzed by RNase L, an endonuclease that becomes activated upon the activation of innate immune response to cleave viral and host mRNAs as one way to limit viral spread (Han et al. 2014; Huang et al. 2014). Structural analysis by an optimized in cell SHAPE-MaP assay (Smola et al. 2015) revealed that endogenous circRNAs tend to form 16- to 26-bp imperfect RNA duplexes and act as inhibitors of a group of nucleic acids receptors with antiviral activities including the IFN-inducible isoform of adenosine deaminase acting on RNA 1 (ADAR1), ADAR1p150, NF90/NF110, and

double-stranded RNA (dsRNA)-activated protein kinase (PKR) (Liu et al. 2019). Compared to other examined proteins, circRNAs exhibit the highest binding preference for PKR and regulate PKR activation. PKR undergoes autophosphorylation and activation by long dsRNAs (>33-bp), but this activation is blocked by short dsRNAs (16- to 33-bp) (Zheng and Bevilacqua 2004). circRNAs, but not their linear cognate mRNAs, are inhibitors of the dsRNA-induced activation of PKR in a sequence-independent, but dsRNA structure–dependent manner. Depleting RNase L in cells resulted in delayed PKR activation, whereas introducing dsRNA-containing circRNAs into cells rendered them susceptible to EMCV infection (Liu et al. 2019). This regulation is physiologically important, and misregulation has been observed to be related to the autoimmune disease systemic lupus erythematosus (SLE). For example, augmented PKR phosphorylation and circRNA reduction were found in peripheral blood mononuclear cells (PBMCs) derived from patients of SLE. Importantly, introducing the dsRNA-containing circRNA, but not their linear cognate mRNAs, into PBMCs or T cells derived from SLE patients attenuated the aberrant PKR activation cascade.

Collectively, by studying processing, structure and degradation of circRNAs, we have discovered that endogenous circRNAs as a group can dampen innate immune responses (Fig. 1E; Li et al. 2017; Liu et al. 2019). These findings are consistent with a recent report that patients with a genetic defect leading to the accumulation of intron lariat–derived RNA circles are more susceptible to viral infections (Zhang et al. 2018). Different from endogenous circRNAs, there are conflicting reports on whether exogenously produced circular RNAs themselves trigger immune responses (Chen et al. 2017; Chen et al. 2019; Wesselhoeft et al. 2019). Future studies are warranted to clarify the modes of action of in vitro–synthesized versus endogenous circular RNAs in the regulation and application in the innate immunity.

IDENTIFICATION OF snoRNA-ENDED LONG NONCODING RNAs

Although it is generally believed that most introns or intron fragments are unstable (Rodriguez-Trelles et al. 2006), the RNA-seq identified a number of nonannotated non-poly(A) RNA signals that mapped to intronic regions in hES H9 and HeLa cells (Yang et al. 2011). Intriguingly, careful analysis of these nonpolyadenylated reads revealed that some excised introns are stabilized by small nucleolar ribonucleoprotein complexes (snoRNPs) at each end, which were named as *sno-lncRNA*s (snoRNA ended long noncoding RNAs) (Fig. 2A). snoRNAs are a family of conserved nuclear RNAs (70–200 nt) that function in the modification of small nuclear RNAs (snRNAs) or processing of ribosomal RNAs (Kiss 2001). Binding of core proteins cotranscriptionally is essential to protect the termini of mature snoRNAs from exonucleolytic degradation (Samarsky et al. 1998; Kufel et al. 2000). In many higher eukaryotes, the great majority of snoRNAs are
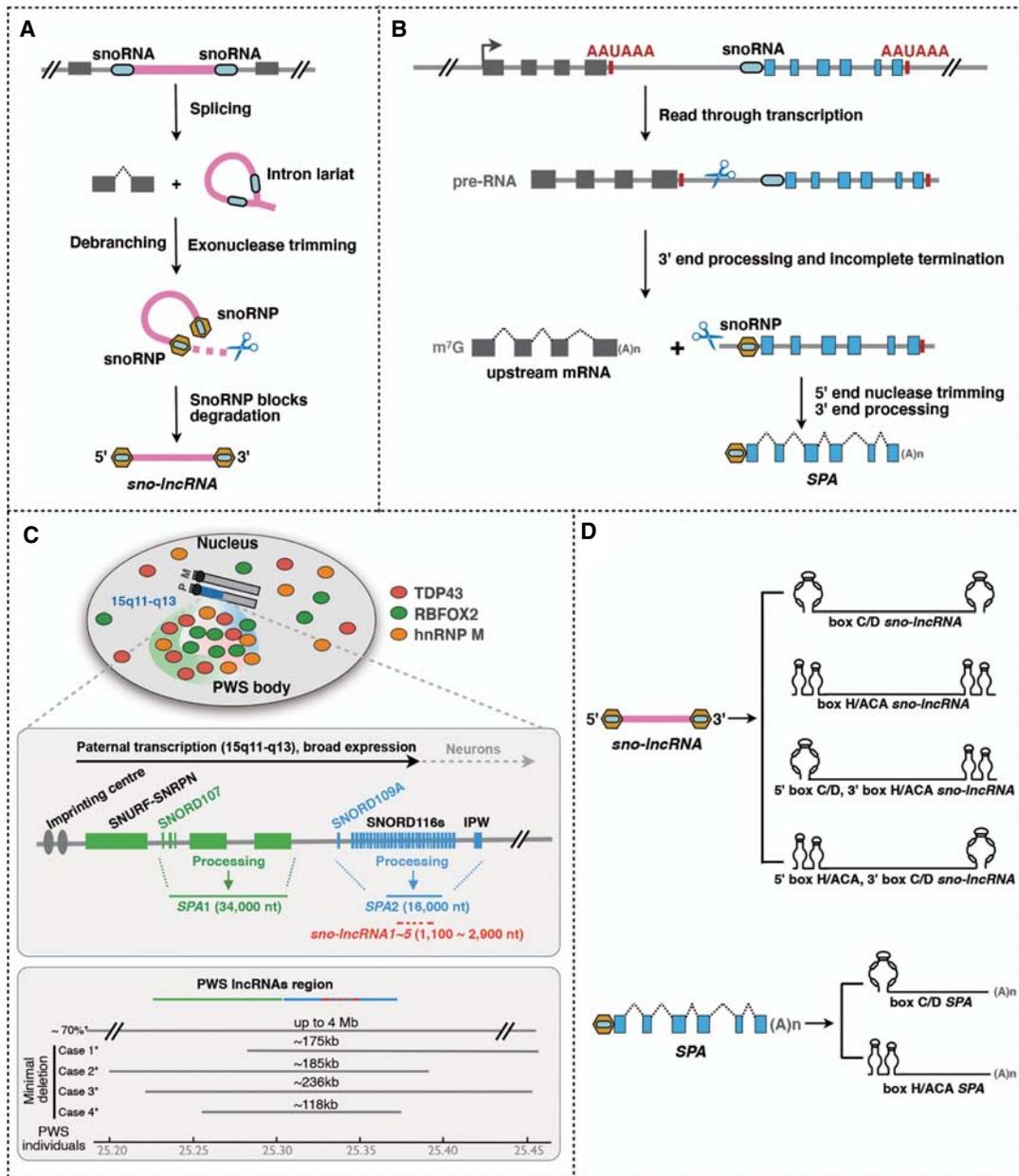
Figure 2. The biogenesis and diversity of snoRNA-ended lncRNAs. (*A*) A model for snoRNA-ended long noncoding RNA (*sno-lncRNA*) processing. Typically, when one intron contains two snoRNA genes, the cotranscriptional formation of snoRNPs at the ends can protect the intronic sequences from exonuclease trimming after debranching during splicing, leading to the formation of sno-lncRNAs. (*B*) A model for 5′-end snoRNP-capped, 3′-polyadenylated (*SPA*) lncRNA processing. Like that of *sno-lncRNA*, *SPA* processing requires an intact snoRNA at the 5′ end. A weak poly(A) signal located downstream from the coding region of mRNA is also necessary for generating readthrough transcripts of pre-*SPAs*. (*C*) snoRNA-ended lncRNAs are derived from the critical region deleted in PWS patients. (*Top*) *sno-lncRNAs* and *SPAs* derived from the PWS deletion region (human 15q11-q13) interact with multiple RBPs including TDP43, RBFOX2, and hnRNP M in the nucleus of human ES cells. (*Bottom*) Minimal chromosome deletions reported in four PWS individuals (Cases 1–4) (Sahoo et al. 2008; de Smith et al. 2009; Duker et al. 2010; Bieth et al. 2015). (*D*) Subtypes of snoRNA-ended lncRNAs. *sno-lncRNAs*, and *SPAs* can be ended with either box C/D or box H/ACA snoRNAs at their ends.

produced from introns, and usually one intron only contains one snoRNA (Weinstein and Steitz 1999; Dieci et al. 2009). When one intron contains two snoRNA genes, the sequences between the snoRNAs are not degraded after splicing, leading to the accumulation of lncRNAs flanked by snoRNA sequences but lacking 5′ caps and 3′ poly(A) tails (Fig. 2A; Yin et al. 2012). Dozens of *sno-lncRNAs*

have been found in mammalian genomes, whose expression is species-specific, because of species-specific alternative splicing that results in single snoRNA or two snoRNAs in one intron (Zhang et al. 2014b; Xing and Chen 2018).

Further application of fibrillarin–RIP and RNA-seq identified 5′-end snoRNP-capped, 3′ polyadenylated

(*SPA*) lncRNAs. *SPA* processing is associated with kinetic competition of the 5′ to 3′ exonuclease XRN2 and Pol II elongation speed downstream from polyadenylation signals in the nucleus. Following cleavage/polyadenylation of an upstream gene, the downstream uncapped *pre-SPA* RNA is trimmed by XRN2 until this exonuclease reaches the cotranscriptionally assembled snoRNP complex (Fig. 2B). More recent study has revealed that the 5′ to 3′ exonuclease activity of the cytoplasmic nonsense-mediated decay pathway mediated by SMG6-mediated endonucleolytic cleavage can also trigger the production of a box C/D snoRD86-containing *SPA*s in the cytoplasm (*c-SPA*s) (Lykke-Andersen et al. 2018).

INVOLVEMENT OF snoRNA-ENDED lncRNAs IN PRADER–WILLI SYNDROME

Among the snoRNA-ended lncRNAs, we found that five *sno-lncRNA*s (sizes are 1000—3000 nt, named *sno-lncRNA1* to *sno-lncRNA5*) and two *SPA*s (35,000 nt for *SPA1* and 16,000 nt for *SPA2*) are conspicuously missing in Prader–Willi syndrome, a neurodevelopmental genetic disorder with elusive molecular causes (Cassidy et al. 2012; Yin et al. 2012; Chamberlain 2013; Wu et al. 2016; Aman et al. 2018). In normal hESCs, these abundant PWS region snoRNA-ended lncRNAs accumulate in *cis* to form one 1- to $\sim$2-μm^3 RNA–protein puncta per nucleus. Importantly, these lncRNAs interact with different RNA binding proteins (RBPs) including TAR DNA-binding protein 43 (TDP43), heterogeneous nuclear ribonucleoprotein M (hnRNPM), and RBP fox-1 homolog 2 (RBFOX2) in normal hESCs and other examined human cell lines. Super resolution microscopy and iCLIP further showed strong preferences between different sno-ended lncRNAs and their interacting RBPs: *SPA1* preferred to interact with TDP43, *sno-lncRNA*s preferred to bind RBFOX2, whereas *SPA2* did not have a preference to these examined RBPs. Generation of a PWS cellular model in hESCs by depleting these lncRNAs using CRISPR–Cas9 revealed the mislocalization of their associated RBPs and an altered pattern of alternative splicing of more than 300 mRNAs without affecting the expression level of individual corresponding mRNAs (Fig. 2C). Importantly, some genes with altered splicing regulation are associated with synaptosome and neurotrophin signaling pathways (Yin et al. 2012; Wu et al. 2016). These studies together indicate that missing of these snoRNA-related lncRNAs is linked to PWS pathogenesis (Fig. 2C).

Interestingly, the lncRNAs *116HG* from the Prader–Willi locus in mouse also form similar cloud-like nuclear accumulations (Powell et al. 2013; Coulson et al. 2018), but currently existing PWS mouse models do not fully recapitulate the PWS patient phenotypes (Ding et al. 2008), highlighting the importance of developing human PWS cellular models in studying the pathogenesis of PWS in the future.

Future work is warranted to examine whether hESCs lacking all these lncRNAs would have any measurable effect on neuronal functions, especially hypothalamic neurons, including not only the morphological phenotypes but also alterations in transcriptomes and alternative splicing patterns. In addition, the contributions of *sno-lncRNA*s and *SPA*s to the molecular and disease phenotype at this locus should also be pursued by identifying additional proteins, DNAs and RNAs that are associated with individual *sno-lncRNA*s and *SPA*s in hESCs and differentiated neurons. Nevertheless, these findings expand the diversity of lncRNAs and provide previously unappreciated insights into PWS pathogenesis.

REGULATION OF RNA POLYMERASE I TRANSCRIPTION BY *SLERT* IN THE HUMAN NUCLEOLUS

There are two main classes of snoRNAs: box C/D snoRNAs and box H/ACA snoRNAs in eukaryotes according to their conserved sequence motifs (Reichow et al. 2007). Thus, four different subtypes of *sno-lncRNA*s that each contain the same or different box C/D or box H/ACA snoRNPs at the ends might be expected to exist (Fig. 2D; Zhang et al. 2014b). In addition to the PWS region *sno-lncRNA*s that are capped by a box C/D snoRNA at each end (Yin et al. 2012), *SLERT* (snoRNA-ended lncRNA enhances pre-ribosomal RNA transcription) is a *sno-lncRNA* that contains box H/ACA snoRNAs at its ends and is highly expressed in multiple human cell lines (Fig. 3A; Xing et al. 2017). Of note, *sno-lncRNA*s containing a box C/D or a box H/ACA snoRNP at each end were also observed in human cells (Zhang et al. 2014b). Similarly, in addition to the 5′ box C/D snoRNA-ended PWS region *SPA*s and *c-SPA*s (Wu et al. 2016; Lykke-Andersen et al. 2018), box H/ACA snoRNA-ended *SPA*s were also detected in several human cell lines (Luo, We, Chen, et al., unpubl. data) (Fig. 2D).

SLERT is generated from the intron of the transforming growth factor beta regulator 4 (*TBRG4*) gene locus and alters RNA polymerase I (Pol I) transcription of ribosomal RNAs (rRNAs) (Xing et al. 2017). Alternative splicing leading to skipping of exons 4 and 5 of the *TBRG4* locus results in two box H/ACA snoRNAs embedded within one intron, which subsequently forms *SLERT* (Fig. 3A). Unlike PWS region *sno-lncRNA*s, *SLERT* does not localize to its own transcription site but instead accumulates in the nucleolus, which is the largest nuclear subdomain in which rRNA biogenesis takes place. Translocation of *SLERT* from its transcription site to the nucleolus depends on its two box H/ACA snoRNAs (Fig. 3B).

Once localized in the nucleolus, *SLERT* directly interacts with the DEAD-box family protein 21 (DDX21) via a 143-nt-long internal region within the two snoRNAs. DDX21 is a DEAD-box RNA helicase that is involved in multiple steps of ribosome biogenesis by contacting both rRNA and snoRNAs and is thought to modulate rRNA transcription, processing, and modification (Holmstrom et al. 2008; Calo et al. 2015; Sloan et al. 2015). Applying super-resolution structured illumination microscopy (SIM) to examine DDX21 localization in live and fixed human cells in detail, we found that DDX21 is
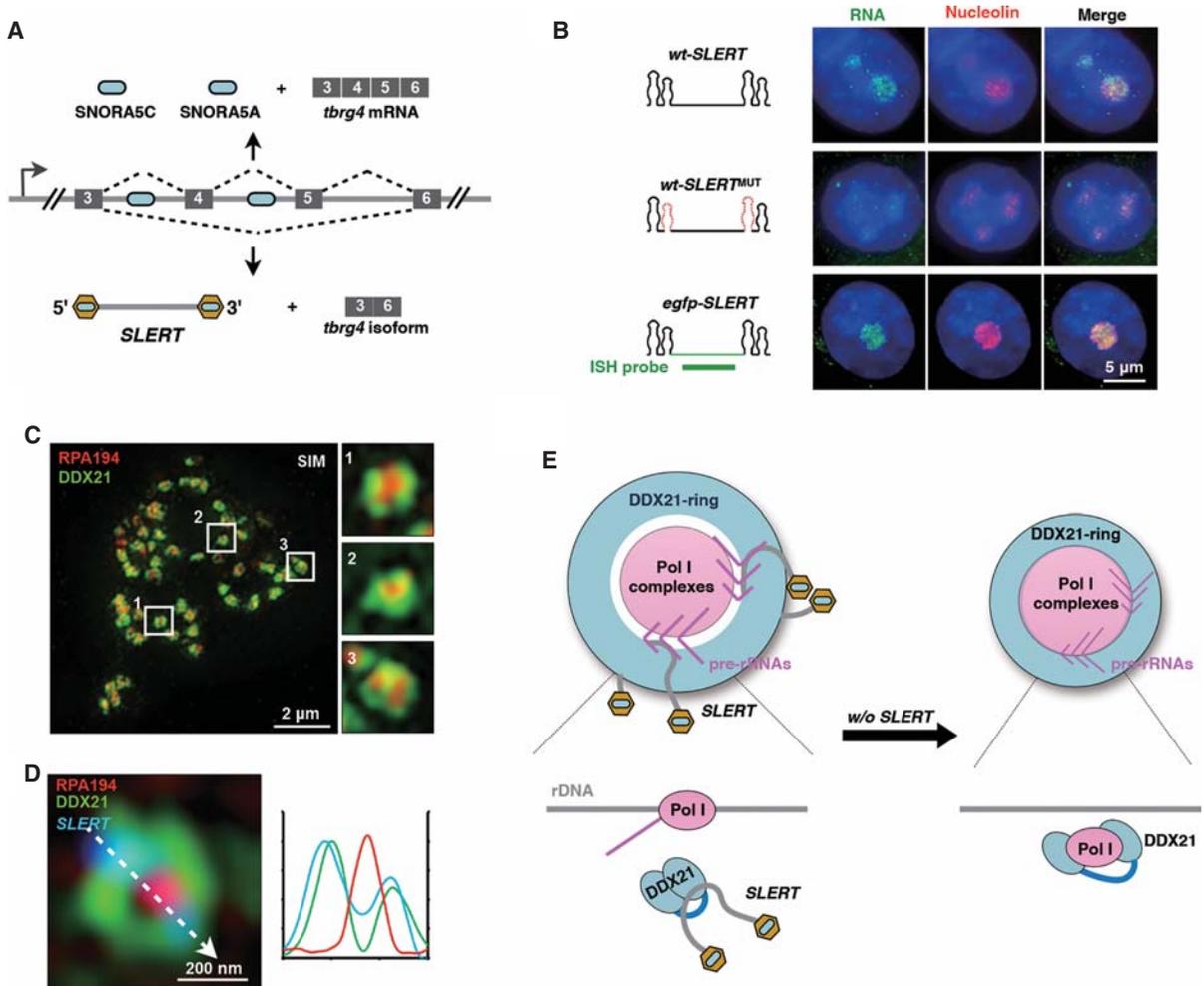
Figure 3. The biogenesis and function of *SLERT*. (*A*) Alternative splicing of the *TBRG4* locus. Alternative splicing of *tbrg4* pre-mRNA generates either two snoRNAs and the *tbrg4* mRNA (*top*) or a snoRNA-ended lncRNA that enhances preribosomal RNA transcription (*SLERT*) and a short isoform of *tbrg4* (rapidly degraded) with skipped exons 4 and 5. (*B*) *SLERT* requires its box H/ACA snoRNA ends to be translocated to the nucleolus. Each indicated *SLERT* or *SLERT* mutants were expressed in HeLa cells, followed by co-staining of WT-*SLERT*, WT-*SLERT*[MUT], or *egfp-SLERT* (green) and nucleolar marker nucleolin (red). (*C*) The Pol I complexes are located within DDX21 rings, shown by SIM in PA1 cells. (*D*) *SLERT* interacts with DDX21 rings but not Pol I complexes. A representative image of co-staining DDX21 (green), RPA194 (red), and *SLERT* (blue) by SIM and a plot profile of the image are shown. (*E*) *SLERT* modulates DDX21 in Pol I transcription regulation. In the nucleolus, DDX21 ring-shaped arrangement surrounds multiple Pol I complexes and inhibits Pol I activity. Binding by *SLERT* allosterically alters individual DDX21 molecules, leading to the reduced interaction between DDX21 and Pol I, which subsequently allows the Pol I complexes to occupy the actively transcribed rDNAs. (*A–E*, Adapted and modified from Xing et al. 2017.)

largely enriched in the nucleolus and form dozens of ring-like structures encircling Pol I complexes (Fig. 3C). Further biochemical analyses showed that DDX21 interacts with subunits of Pol I complexes and prevents them loading onto rDNAs in an RNA helicase activity–independent manner and subsequently inhibits pre-rRNA transcription.

Depleting one snoRNA of *SLERT* by CRSIPR–Cas9 abolished *SLERT* expression in examined human cells and led to suppressed Pol I transcription. Under SIM observation, *SLERT* does not directly interact with Pol I, but instead specifically localizes to individual DDX21 rings (Fig. 3D). Mechanistically, *SLERT* binding to DDX21 alters the conformation of individual DDX21 molecules, suppressing the interaction between DDX21 and Pol I,

therefore releasing Pol I to be engaged to rDNAs for active pre-RNA transcription (Fig. 3E).

ULTRA-STRUCTURE ORGANIZATION OF THE HUMAN NUCLEOLUS

The primary function of the nucleolus serves as the site of rRNA biogenesis. It has been well-established that the mammalian nucleolus is highly organized and is comprised of three morphologically distinct subregions and named fibrillar centers (FCs), dense fibrillar components (DFCs), and the granular component (GC), shown by the electron density under electron microscopy (EM) (Fig. 4A; Koberna et al. 2002; Boisvert et al. 2007). Recent single-molecule images have further revealed the tripartite
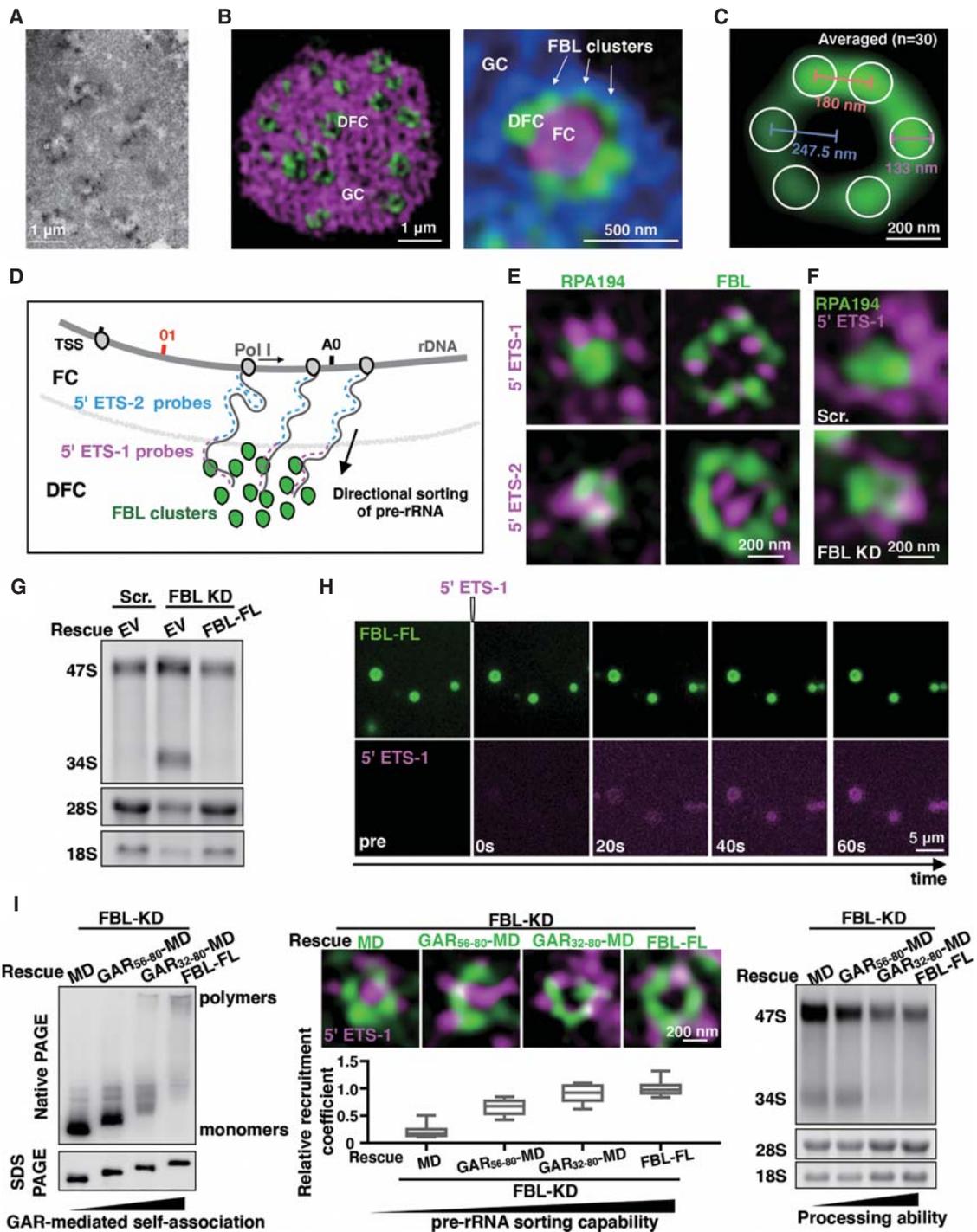
Figure 4. Ultra-structure organization of the human nucleolus and nascent pre-rRNA translocation is required for proper pre-rRNA processing and function. (*A*) Electron microscopy shows the localization of nucleolus in thin-sectioned permeabilized HeLa cells. (*B*) Representative SIM images of nucleoli and three nucleolar subregions in live HeLa cells. (*C*) Cross-correlation of aligned and averaged images shows that the max-cross sections of individual DFC regions contain six FBL clusters. The minimum distance between two clusters is ~180 nm; the diameter of FBL cluster is ~133 nm; the distance between the center of individual clusters and the center of DFC is ~247.5 nm; the detached distance between two adjacent PF clusters is ~180 nm. (*D*) A schematic of RNA smFISH probes to detect transcribing pre-rRNA. The 5′ ETS-1 probe detects nt 1–414 of pre-rRNAs; the 5′ ETS-2 probe detects nt 498–977 of pre-rRNA. (*E*) The 5′ ETS-1 probe-detects 47S pre-rRNAs that are largely distributed outside of the FC, whereas the 5′ ETS-2 probe detects pre-rRNAs that are mainly located at the FC/DFC border, shown by SIM. (*F*) FBL knockdown (KD) results in impaired localization of nascent pre-rRNAs shown by the highest colocalization signal between RPA194 and 5′ ETS-1. (*G*) FBL KD results in aberrant accumulation of 47S and 34S pre-rRNAs, accompanied by reduced 28S and 18S rRNAs. (*H*) Cy3-labeled 5′ ETS-1 (magenta) is sorted to the mNeonGreen-FBL-FL droplets (green) in vitro. (*I*) The IDR length in the GAR domain of FBL promotes FBL self-association, which confers the capability of pre-rRNA sorting and processing in which FBL is involved. (*Left*) Increased GAR domain length in FBL mutants leads to augmented FBL self-association shown by increased multimerization. (*Middle*) Increased IDR length in the GAR domain of FBL mutants promotes translocation of pre-rRNAs into the DFC region. Twenty cells were analyzed under each condition by boxplot. (*Right*) Increased IDR length in the GAR domain of FBL positively correlates with proper 47S pre-rRNA processing, shown by northern blots. (*A*, Adapted from Koberna et al. 2002; *B–I*, adapted and modified from Yao et al. 2019b.)

nucleolar organization (Szczurek et al. 2016; Khan et al. 2018). Continuous Pol I transcription at the FC produces nascent pre-rRNAs and causes the subsequent radial flux of pre-rRNAs through the DFC for processing and modification to produce 28S and 18S rRNAs, and then into the GC for ribosome assembly and finally into the nucleoplasm (Boisvert et al. 2007; McStay and Grummt 2008). A human nucleolus is assembled around active nucleolar organizer regions (NORs), which are composed of clusters of tandem repeats of ribosomal DNA (rDNA), each with a long intergenic spacer (IGS) of ~30 kb and a pre-ribosomal RNA (pre-rRNA) coding region of ~14 kb (McStay and Grummt 2008).

The SIM observation that DDX21 forms dozens of ring-like structures surrounding Pol I complexes (Fig. 3C) raised an intriguing question—how are DDX21-rings positioned in the tripartite organization of the mammalian nucleolus? Other questions remaining to be addressed include where does Pol I transcription occur (Mais and Scheer 2001; Cheutin et al. 2002; Huang 2002; Boisvert et al. 2007) and how do nascent pre-rRNAs migrate in the nucleolus? A time-dependent migration of Br-U-labeled pre-rRNAs from FCs to DFCs was observed in HeLa cells (Thiry et al. 2000), but the underlying mechanism of this observation had remained unclear.

We applied CRISPR–Cas9-mediated knock-in of fluorescently tagged proteins to visualize FC, DFC, and GC subnucleolar organization under SIM that allowed us to uncover previously uncharacterized nucleolar ultra-structures (Fig. 4B; Yao et al. 2019b). A human nucleolus consists of dozens of FC/DFC units that are assembled around two to three copies of active rDNAs at the border of each FC/DFC, where Pol I complexes are located and Pol I transcription occurs. Pre-rRNA processing factors, such as fibrillarin (FBL) form 18 to 24 clusters that are further assembled into a polyhedron-like shell of the DFC surrounding FC. On average, each spherical PF cluster is ~133 nm in diameter and each DFC region contains ~628 nm outer and ~362 nm inner diameters (Fig. 4C). Consistent with this model, Pol I complex subunits are enriched at the FC border as clusters (Yao et al. 2019b), which are in striking contrast to the previous models in which Pol I complexes were thought to be distributed throughout the entire FC region (Cheutin et al. 2002). Further, an active NOR contains not only active rDNAs but also transcriptionally inert rDNAs, consistent with the earlier observation of discontinuous transcribed rDNA clusters in rDNA spreads (McKnight and Miller 1976; McStay and Grummt 2008). Collectively, these observations represent a substantial advance in our understanding of nucleolar spatial organization.

NASCENT pre-rRNA TRANSLOCATION IS REQUIRED FOR PROPER pre-rRNA PROCESSING AND FUNCTION

An increasing number of RNAs have emerged as important modulators of nuclear structure and function by acting in *cis* or in *trans* (Bergmann and Spector 2014;

Chen 2016; Engreitz et al. 2016; Tomita et al. 2017; Yao et al. 2019a). What mechanisms does the cell use to keep nascent RNAs from sticking together while also promoting directional sorting in *trans*? It has been a challenge to address this question in single cells because of the low abundance of most nascent RNAs that usually undergo rapid processing. Remarkably, we observed that the 5′ termini of the nascent 47S pre-rRNAs are translocated to the DFCs, whereas pre-rRNAs are still being transcribed at the border of FC and DFC (Fig. 4D,E), indicating that the relatively high abundance of nascent 47S pre-rRNA and its radial flux mode of processing in FC/DFCs can be an attractive model to study how nascent RNA directional sorting is achieved.

Earlier studies in yeast revealed that the eukaryotic pre-rRNA processing is complex. The early step was thought to be involved in the assembly of small subunit (SSU) processomes including UTPa, UTPb, and U3 snoRNP (Barandun et al. 2018). UTPa complexes were reported to bind first to nascent 35S pre-rRNA to chaperone pre-rRNA and U3 snoRNA to initiate SSU assembly (Hunziker et al. 2016). Such an initial binding between UTPa and the 5′ ETS appeared to be required for the subsequent recruitment of the UTPb complex and the U3 snoRNP components (Henras et al. 2015; Sharma and Lafontaine 2015; Barandun et al. 2018). These studies in yeast might not directly reflect processing of nascent pre-rRNA in the human nucleolus since yeast cells have bipartite nucleoli containing merged FC/DFCs and GC (Thiry and Lafontaine 2005), which is distinct from tripartite nucleoli in human cells, where pre-rRNA processing takes place in DFCs.

Use of shRNA-mediated knockdown of factors in UTPa/b and snoRNP complexes combined with structured illumination microscopy (SIM)-based screening have revealed that FBL plays a key role in promoting the movement of the 5′ end of 47S pre-rRNA from its transcription site to DFC. Depletion of FBL dramatically blocks the translocation of 5′ ends of transcribing pre-rRNAs and results in the accumulation of pre-rRNAs at the transcription sites at the border of FC/DFCs (Fig. 4F), together with aberrant pre-rRNA processing shown as 34S pre-rRNA accumulation (Fig. 4G).

FBL is well-known as an rRNA 2′-O-methyltransferase in the snoRNP particles that participate in pre-rRNA processing (Tollervey et al. 1991; Tollervey et al. 1993; Tafforeau et al. 2013) and Pol I transcription (Tessarz et al. 2014; Loza-Muller et al. 2015). Its amino-terminus contains a glycine- and arginine-rich (GAR) domain highly enriched with intrinsically disordered regions (IDRs) that are required for *Xenopus laevis* FBL phase separation (Feric et al. 2016); its carboxyl terminus contains an RNA binding domain (RBD) and the α domain for methyltransferase activity (MD domain).

How does FBL control the movement of the 5′ end of 47S pre-rRNA to the DFC? First, distinct from other components in snoRNP complexes, FBL specifically and strongly interacts with the 5′ end of 47S pre-rRNA as revealed by PAR-CLIP (Kishore et al. 2013) and by in vitro binding and phase-separation assays (Yao et al.

2019b). These observations indicate an additional function of FBL in nascent pre-rRNA sorting beyond its classical role in U3 snoRNP. Second, the GAR domain is necessary and sufficient for human FBL self-aggregation into droplets at physiological concentrations (Yao et al. 2019b). Importantly, both in vitro and in vivo lines of evidence suggest a model whereby the binding and sorting of nascent 47S pre-rRNA utilizes different FBL domains: The MD domain binds to the 5′ end of 47S pre-rRNA, and the RNA-interacting FBL then moves toward the DFC by GAR domain self-association (Fig. 4H–I). Such pre-rRNA sorting strongly correlates with FBL self-association via the strength of IDRs in FBL and is required for proper pre-rRNA processing in cells (Fig. 4I). For example, rescue of FBL KD in cells with FBL mutants containing different length of GARs but with intact MD domain shows the enhanced formation of FBL polymers, enhanced 5′-end pre-rRNA sorting to DFC, and increased proper production of 47S pre-rRNA (Fig. 4I). Importantly, this sorting and translocation process is required for proper ribosome production, thus representing yet another example illustrating the importance of linking RNA processing to function.

More broadly, as many RBPs associated with nascent pre-mRNA processing events are known to contain structurally disordered regions (Banani et al. 2017; Gueroussov et al. 2017; Shin and Brangwynne 2017; Ying et al. 2017), a similar liquid–liquid phase-separation controlled nascent rRNA sorting mechanism is likely used by the cell to keep other types of nascent RNAs from unnecessary or unwanted self-aggregation.

CONCLUSION

It has been well-established that Pol II transcription, nascent pre-mRNA splicing, capping, polyadenylation, mRNA export, and surveillance are seamlessly integrated (Moore and Proudfoot 2009). Studies of unconventionally processed lncRNAs including circular RNAs (Fig. 1), snoRNA-related RNAs (Figs. 2 and 3), and the "housekeeping" pre-rRNAs (Fig. 4) have revealed no exception for ncRNAs: Throughout the maturation process of each district class of RNA, transcription and processing are crucially important for subcellular localization and function.

One may argue that all examples illustrated so far are unconventionally processed ncRNAs. Indeed, despite these examples, it has been well-studied that Pol II transcription and the 3′-end alternative processing of the NEAT1 lncRNA act together to modulate paraspeckle morphology and function (Mao et al. 2011; Naganuma et al. 2012; Hirose et al. 2014; Wang et al. 2018; Yamazaki et al. 2018). Emerging studies have also revealed both cis- and trans-factors are required for the subcellular localization of Pol II–transcribed mRNA-like lincRNAs and functions (Hacisuleyman et al. 2014; Zhang et al. 2014b; Lubelsky and Ulitsky 2018; Shukla et al. 2018).

Recent studies using the most robustly available methods have greatly advanced our understanding of cellular

functions of lncRNAs (Kopp and Mendell 2018; Yao et al. 2019a). Beyond linking RNA processing and function, how lncRNAs are structured and what structural conformations lncRNAs adopt for their interacting partners have remained mysteries. Their large sizes and flexible nature have endowed them with previously underappreciated functional potentials; however, this has also presented unexpected experimental challenges to understanding their regulation at multiple levels. Future studies aimed at understanding in greater detail the regulation of expression, subcellular localization patterns, interaction partners, and conformational information of lncRNAs, as well as understanding how the biogenesis and turnover of RNAs are linked to their subcellular localization, will provide greater insight into their cellular functions.

REFERENCES

Adriaens C, Standaert L, Barra J, Latil M, Verfaillie A, Kalev P, Boeckx B, Wijnhoven PW, Radaelli E, Vermi W, et al. 2016. p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. Nat Med 22: 861–868. doi:10.1038/nm.4135

Aman LCS, Manning KE, Whittington JE, Holland AJ. 2018. Mechanistic insights into the genetics of affective psychosis from Prader–Willi syndrome. Lancet Psychiatry 5: 370–378. doi:10.1016/S2215-0366(18)30009-9

Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, Hearn S, Kim Y, MacLeod AR, Krainer AR, Norton L, et al. 2016. Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. Genes Dev 30: 34–51. doi:10.1101/gad.270959.115

Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. 2014. circRNA biogenesis competes with pre-mRNA splicing. Mol Cell 56: 55–66. doi:10.1016/j.molcel.2014.08.019

Banani SF, Lee HO, Hyman AA, Rosen MK. 2017. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18: 285–298. doi:10.1038/nrm.2017.7

Barandun J, Hunziker M, Klinge S. 2018. Assembly and structure of the SSU processome—a nucleolar precursor of the small ribosomal subunit. Curr Opin Struct Biol 49: 85–93. doi:10.1016/j.sbi.2018.01.008

Bartel DP. 2018. Metazoan microRNAs. Cell 173: 20–51. doi:10.1016/j.cell.2018.03.006

Bergmann JH, Spector DL. 2014. Long non-coding RNAs: modulators of nuclear structure and function. Curr Opin Cell Biol 26: 10–18. doi:10.1016/j.ceb.2013.08.005

Bieth E, Eddiry S, Gaston V, Lorenzini F, Buffet A, Conte Auriol F, Molinas C, Cailley D, Rooryck C, Arveiler B, et al. 2015. Highly restricted deletion of the SNORD116 region is implicated in Prader–Willi syndrome. Eur J Hum Genet 23: 252–255. doi:10.1038/ejhg.2014.103

Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. 2007. The multifunctional nucleolus. Nat Rev Mol Cell Biol 8: 574–585. doi:10.1038/nrm2184

Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. 2012. Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. Proc Natl Acad Sci 109: 19202–19207. doi:10.1073/pnas.1217338109

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25: 1915–1927. doi:10.1101/gad.17446611

Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A. 2015. Localization and abundance analysis of human lncRNAs at single-cell and sin-

gle-molecule resolution. *Genome Biol* **16:** 20. doi:10.1186/s13059-015-0586-4

Calo E, Flynn RA, Martin L, Spitale RC, Chang HY, Wysocka J. 2015. RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518:** 249–253. doi:10.1038/nature13923

Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, Goodfellow P, Lovellbadge R. 1993. Circular transcripts of the testis-determining gene *Sry* in adult-mouse testis. *Cell* **73:** 1019–1030. doi:10.1016/0092-8674(93)90279-Y

Carlevaro-Fita J, Johnson R. 2019. Global positioning system: understanding long noncoding RNAs through subcellular localization. *Mol Cell* **73:** 869–883. doi:10.1016/j.molcel.2019.02.008

Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. 2012. Prader–Willi syndrome. *Genet Med* **14:** 10–26. doi:10.1038/gim.0b013e31822bead0

Chamberlain SJ. 2013. RNAs of the human chromosome 15q11-q13 imprinted region. *Wiley Interdiscip Rev RNA* **4:** 155–166. doi:10.1002/wrna.1150

Chen LL. 2016. Linking long noncoding RNA localization and function. *Trends Biochem Sci* **41:** 761–772. doi:10.1016/j.tibs.2016.07.003

Chen CK, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, Cerase A, McDonel P, Guttman M. 2016. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354:** 468–472. doi:10.1126/science.aae0047

Chen YG, Kim MV, Chen XQ, Batista PJ, Aoyama S, Wilusz JE, Iwasaki A, Chang HY. 2017. Sensing self and foreign circular RNAs by intron identity. *Mol Cell* **67:** 228–238. doi:10.1016/j.molcel.2017.05.022

Chen YG, Chen R, Ahmad S, Verma R, Kasturi SP, Amaya L, Broughton JP, Kim J, Cadena C, Pulendran B, et al. 2019. N^6-methyladenosine modification controls circular RNA immunity. *Mol Cell* **76:** 96–109. doi:10.1016/j.molcel.2019.07.016

Cheutin T, O'Donohue MF, Beorchia A, Vandelaer M, Kaplan H, Defever B, Ploton D, Thiry M. 2002. Three-dimensional organization of active rRNA genes within the nucleolus. *J Cell Sci* **115:** 3297–3307.

Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY. 2015. Systematic discovery of Xist RNA binding proteins. *Cell* **161:** 404–416. doi:10.1016/j.cell.2015.03.025

Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33:** 717–726. doi:10.1016/j.molcel.2009.01.026

Cocquerelle C, Daubersies P, Majerus MA, Kerckaert JP, Bailleul B. 1992. Splicing with inverted order of exons occurs proximal to large introns. *EMBO J* **11:** 1095–1098. doi:10.1002/j.1460-2075.1992.tb05148.x

Conn VM, Hugouvieux V, Nayak A, Conos SA, Capovilla G, Cildir G, Jourdain A, Tergaonkar V, Schmid M, Zubieta C, et al. 2017. A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. *Nat Plants* **3:** 17053. doi:10.1038/nplants.2017.53

Coulson RL, Powell WT, Yasui DH, Dileep G, Resnick J, LaSalle JM. 2018. Prader–Willi locus *Snord116* RNA processing requires an active endogenous allele and neuron-specific splicing by *Rbfox3*/NeuN. *Hum Mol Genet* **27:** 4051–4060. doi:10.1093/hmg/ddy296

Creamer KM, Lawrence JB. 2017. XIST RNA: a window into the broader role of RNA in nuclear chromosome architecture. *Philos Trans R Soc Lond B Biol Sci* **372:** 20160360. doi:10.1098/rstb.2016.0360.

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, et al. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22:** 1775–1789. doi:10.1101/gr.132159.111

de Smith AJ, Purmann C, Walters RG, Ellis RJ, Holder SE, Van Haelst MM, Brady AF, Fairbrother UL, Dattani M, Keogh JM, et al. 2009. A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. *Hum Mol Genet* **18:** 3257–3265. doi:10.1093/hmg/ddp263

Dhir A, Dhir S, Proudfoot NJ, Jopling CL. 2015. Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs. *Nat Struct Mol Biol* **22:** 319–327. doi:10.1038/nsmb.2982

Dieci G, Preti M, Montanini B. 2009. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* **94:** 83–88. doi:10.1016/j.ygeno.2009.05.002

Ding F, Li HH, Zhang S, Solomon NM, Camper SA, Cohen P, Francke U. 2008. SnoRNA *Snord116* (*Pwcr1/MBII-85*) deletion causes growth deficiency and hyperphagia in mice. *PLoS One* **3:** e1709. doi:10.1371/journal.pone.0001709

Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, et al. 2012. Landscape of transcription in human cells. *Nature* **489:** 101–108. doi:10.1038/nature11233

Dong R, Ma XK, Chen LL, Yang L. 2017. Increased complexity of circRNA expression during species evolution. *RNA Biol* **14:** 1064–1074. doi:10.1080/15476286.2016.1269999

Duker AL, Ballif BC, Bawle EV, Person RE, Mahadevan S, Alliman S, Thompson R, Traylor R, Bejjani BA, Shaffer LG, et al. 2010. Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader–Willi syndrome. *Eur J Hum Genet* **18:** 1196–1201. doi:10.1038/ejhg.2010.102

Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, et al. 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341:** 1237973. doi:10.1126/science.1237973

Engreitz JM, Ollikainen N, Guttman M. 2016. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat Rev Mol Cell Biol* **17:** 756–770. doi:10.1038/nrm.2016.126

Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP. 2016. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165:** 1686–1697. doi:10.1016/j.cell.2016.04.047

Gardner EJ, Nizami ZF, Talbot CC, Gall JG. 2012. Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. *Genes Dev* **26:** 2550–2559. doi:10.1101/gad.202184.112

Goff LA, Rinn JL. 2015. Linking RNA biology to lncRNAs. *Genome Res* **25:** 1456–1465. doi:10.1101/gr.191122.115

Gueroussov S, Weatheritt RJ, O'Hanlon D, Lin ZY, Narula A, Gingras AC, Blencowe BJ. 2017. Regulatory expansion in mammals of multivalent hnRNP assemblies that globally control alternative splicing. *Cell* **170:** 324–339.e323. doi:10.1016/j.cell.2017.06.037

Guo JU, Agarwal V, Guo HL, Bartel DP. 2014. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* **15:** 409. doi:10.1186/s13059-014-0409-z

Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson DG, Sauvageau M, Kelley DR, et al. 2014. Topological organization of multi-chromosomal regions by the long intergenic noncoding RNA *Firre*. *Nat Struct Mol Biol* **21:** 198–206. doi:10.1038/nsmb.2764

Han YC, Donovan J, Rath S, Whitney G, Chitrakar A, Korennykh A. 2014. Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. *Science* **343:** 1244–1248. doi:10.1126/science.1249845

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* **495:** 384–388. doi:10.1038/nature11993

Harashima A, Guettouche T, Barber GN. 2010. Phosphorylation of the NFAR proteins by the dsRNA- dependent protein kinase

PKR constitutes a novel mechanism of translational regulation and cellular defense. *Genes Dev* **24:** 2640–2653. doi:10.1101/gad.1965010

Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A, Gleizes PE. 2015. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA* **6:** 225–242. doi:10.1002/wrna.1269

Hirose T, Virnicchi G, Tanigawa A, Naganuma T, Li R, Kimura H, Yokoi T, Nakagawa S, Benard M, Fox AH, et al. 2014. NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol Biol Cell* **25:** 169–183. doi:10.1091/mbc.e13-09-0558

Holmstrom TH, Mialon A, Kallio M, Nymalm Y, Mannermaa L, Holm T, Johansson H, Black E, Gillespie D, Salminen TA, et al. 2008. c-Jun supports ribosomal RNA processing and nucleolar localization of RNA helicase DDX21. *J Biol Chem* **283:** 7046–7053. doi:10.1074/jbc.M709613200

Hsu MT, Cocaprados M. 1979. Electron-microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* **280:** 339–340. doi:10.1038/280339a0

Huang S. 2002. Building an efficient factory: where is pre-rRNA synthesized in the nucleolus? *J Cell Biol* **157:** 739–741. doi:10.1083/jcb.200204159

Huang H, Zeqiraj E, Dong BH, Jha BK, Duffy NM, Orlicky S, Thevakumaran N, Talukdar M, Pillon MC, Ceccarelli DF, et al. 2014. Dimeric structure of pseudokinase RNase L bound to 2–5 Å reveals a basis for interferon-induced antiviral activity. *Mol Cell* **53:** 221–234. doi:10.1016/j.molcel.2013.12.025

Hunziker M, Barandun J, Petfalski E, Tan D, Delan-Forino C, Molloy KR, Kim KH, Dunn-Davies H, Shi Y, Chaker-Margot M, et al. 2016. UtpA and UtpB chaperone nascent pre-ribosomal RNA and U3 snoRNA to initiate eukaryotic ribosome assembly. *Nat Commun* **7:** 12090. doi:10.1038/ncomms12090

Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8:** 39. doi:10.1186/1471-2164-8-39

Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, et al. 2015. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep* **10:** 170–177. doi:10.1016/j.celrep.2014.12.019

Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu JZ, Marzluff WF, Sharpless NE. 2013. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19:** 141–157. doi:10.1261/rna.035667.112

Jia Ng SS, Zheng RT, Osman I, Pek JW. 2018. Generation of *Drosophila* sisRNAs by independent transcription from cognate introns. *iScience* **4:** 68–75. doi:10.1016/j.isci.2018.05.010

Khan S, Verma NC, Chethana, Nandi CK. 2018. Carbon dots for single-molecule imaging of the nucleolus. *ACS Applied Nano Mater* **1:** 483–487. doi:10.1021/acsanm.7b00175

Kishore S, Gruber AR, Jedlinski DJ, Syed AP, Jorjani H, Zavolan M. 2013. Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol* **14:** R45. doi:10.1186/gb-2013-14-5-r45

Kiss T. 2001. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J* **20:** 3617–3622. doi:10.1093/emboj/20.14.3617

Koberna K, Malinsky J, Pliss A, Masata M, Vecerova J, Fialova M, Bednar J, Raska I. 2002. Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ. *J Cell Biol* **157:** 743–748. doi:10.1083/jcb.200202007

Kopp F, Mendell JT. 2018. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172:** 393–407. doi:10.1016/j.cell.2018.01.011

Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. 2019. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet* **20:** 675–691. doi:10.1038/s41576-019-0158-7

Kufel J, Allmang C, Chanfreau G, Petfalski E, Lafontaine DL, Tollervey D. 2000. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol* **20:** 5415–5424. doi:10.1128/MCB.20.15.5415-5424.2000

Lee S, Kopp F, Chang TC, Sataluri A, Chen B, Sivakumar S, Yu H, Xie Y, Mendell JT. 2016. Noncoding RNA *NORAD* regulates genomic stability by sequestering PUMILIO proteins. *Cell* **164:** 69–80. doi:10.1016/j.cell.2015.12.017

Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, et al. 2017. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol Cell* **66:** 22–37. doi:10.1016/j.molcel.2017.02.017

Li ZY, Huang C, Bao C, Chen L, Lin M, Wang XL, Zhong GL, Yu B, Hu WC, Dai LM, et al. 2015. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* **22:** 256–264. doi:10.1038/nsmb.2959

Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, Wei J, Yao RW, Yang L, Chen LL. 2017. Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. *Mol Cell* **67:** 214–227. doi:10.1016/j.molcel.2017.05.023

Li X, Yang L, Chen LL. 2018. The biogenesis, functions, and challenges of circular RNAs. *Mol Cell* **71:** 428–442. doi:10.1016/j.molcel.2018.06.034

Liang D, Tatomer DC, Luo Z, Wu H, Yang L, Chen LL, Cherry S, Wilusz JE. 2017. The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. *Mol Cell* **68:** 940–954. doi:10.1016/j.molcel.2017.10.034

Liu CX, Li X, Nan F, Jiang S, Gao X, Guo SK, Xue W, Cui YG, Dong KG, Ding HH, et al. 2019. Structure and degradation of circular RNAs regulate PKR activation in innate immunity. *Cell* **177:** 865–880. doi:10.1016/j.cell.2019.03.046

Loza-Muller L, Rodríguez-Corona U, Sobol M, Rodríguez-Zapata LC, Hozak P, Castano E. 2015. Fibrillarin methylates H2A in RNA polymerase I trans-active promoters in *Brassica oleracea*. *Front Plant Sci* **6:** 976. doi:10.3389/fpls.2015.00976

Lu TT, Cui LL, Zhou Y, Zhu CR, Fan DL, Gong H, Zhao Q, Zhou CC, Zhao Y, Lu DF, et al. 2015. Transcriptome-wide investigation of circular RNAs in rice. *RNA* **21:** 2076–2087. doi:10.1261/rna.052282.115

Lubelsky Y, Ulitsky I. 2018. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **555:** 107–111. doi:10.1038/nature25757

Lykke-Andersen S, Ardal BK, Hollensen AK, Damgaard CK, Jensen TH. 2018. Box C/D snoRNP autoregulation by a *cis*-acting snoRNA in the *NOP56* pre-mRNA. *Mol Cell* **72:** 99–111.e115. doi:10.1016/j.molcel.2018.08.017

Mais C, Scheer U. 2001. Molecular architecture of the amplified nucleoli of *Xenopus* oocytes. *J Cell Sci* **114:** 709–718.

Malakar P, Shilo A, Mogilevsky A, Stein I, Pikarsky E, Nevo Y, Benyamini H, Elgavish S, Zong X, Prasanth KV, et al. 2017. Long noncoding RNA MALAT1 promotes hepatocellular carcinoma development by SRSF1 upregulation and mTOR activation. *Cancer Res* **77:** 1155–1167. doi:10.1158/0008-5472.CAN-16-1508

Mao YS, Sunwoo H, Zhang B, Spector DL. 2011. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* **13:** 95–101. doi:10.1038/ncb2140

McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521:** 232–236. doi:10.1038/nature14443

McKnight SL, Miller OL Jr. 1976. Ultrastructural patterns of RNA synthesis during early embryogenesis of *Drosophila melanogaster*. *Cell* **8:** 305–319. doi:10.1016/0092-8674(76)90014-3

McStay B, Grummt I. 2008. The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu Rev Cell Dev Biol* **24:** 131–157. doi:10.1146/annurev.cellbio.24.110707.175259

Mele M, Mattioli K, Mallard W, Shechner DM, Gerhardinger C, Rinn JL. 2017. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. *Genome Res* **27:** 27–37. doi:10.1101/gr.214205.116

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, et al. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495:** 333–338. doi:10.1038/nature11928

Minajigi A, Froberg J, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W, et al. 2015. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349:** aab2276. doi:10.1126/science.aab2276.

Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. 2010. Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* **330:** 1244–1247. doi:10.1126/science.1195858

Moore MJ, Proudfoot NJ. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136:** 688–700. doi:10.1016/j.cell.2009.02.001

Morgan JT, Fink GR, Bartel DP. 2019. Excised linear introns regulate growth in yeast. *Nature* **565:** 606–611. doi:10.1038/s41586-018-0828-1

Mudge JM, Frankish A, Harrow J. 2013. Functional transcriptomics in the post-ENCODE era. *Genome Res* **23:** 1961–1973. doi:10.1101/gr.161315.113

Munschauer M, Nguyen CT, Sirokman K, Hartigan CR, Hogstrom L, Engreitz JM, Ulirsch JC, Fulco CP, Subramanian V, Chen J, et al. 2018. The *NORAD* lncRNA assembles a topoisomerase complex critical for genome stability. *Nature* **561:** 132–136. doi:10.1038/s41586-018-0453-z

Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. 2012. Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* **31:** 4020–4034. doi:10.1038/emboj.2012.251

Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, Prasanth KV. 2012. Malat1 is not an essential component of nuclear speckles in mice. *RNA* **18:** 1487–1499. doi:10.1261/rna.033217.112

Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, Kinzler KW, Vogelstein B. 1991. Scrambled exons. *Cell* **64:** 607–613. doi:10.1016/0092-8674(91)90244-S

Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, et al. 2017. Translation of circRNAs. *Mol Cell* **66:** 9–21. doi:10.1016/j.molcel.2017.02.021

Parenteau J, Maignon L, Berthoumieux M, Catala M, Gagnon V, Abou Elela S. 2019. Introns are mediators of cell response to starvation. *Nature* **565:** 612–617. doi:10.1038/s41586-018-0859-7

Piwecka M, Glazar P, Hernandez-Miranda LR, Memczak S, Wolf SA, Rybak-Wolf A, Filipchyk A, Klironomos F, Jara CAC, Fenske P, et al. 2017. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science* **357:** eaam8526. doi:10.1126/science.aam8526

Powell WT, Coulson RL, Crary FK, Wong SS, Ach RA, Tsang P, Alice Yamada N, Yasui DH, Lasalle JM. 2013. A Prader–Willi locus lncRNA cloud modulates diurnal genes and energy expenditure. *Hum Mol Genet* **22:** 4318–4328. doi:10.1093/hmg/ddt281

Quinn JJ, Chang HY. 2016. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17:** 47–62. doi:10.1038/nrg.2015.10

Reichow SL, Hamma T, Ferre-D'Amare AR, Varani G. 2007. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* **35:** 1452–1464. doi:10.1093/nar/gkl1172

Rodriguez-Trelles F, Tarrio R, Ayala FJ. 2006. Origins and evolution of spliceosomal introns. *Annu Rev Genet* **40:** 47–76. doi:10.1146/annurev.genet.40.110405.090625

Rybak-Wolf A, Stottmeister C, Glazar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, et al. 2015. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol Cell* **58:** 870–885. doi:10.1016/j.molcel.2015.03.027

Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, Garnica A, Cheung SW, Beaudet AL. 2008. Prader–Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet* **40:** 719–721. doi:10.1038/ng.158

Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. 2012. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* **7:** e30733. doi:10.1371/journal.pone.0030733

Salzman J, Chen RE, Olsen MN, Wang PL, Brown PO. 2013. Cell-type specific features of circular rna expression. *PLoS Genet* **9:** e1003777. doi:10.1371/journal.pgen.1003777

Samarsky DA, Fournier MJ, Singer RH, Bertrand E. 1998. The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J* **17:** 3747–3757. doi:10.1093/emboj/17.13.3747

Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK. 1976. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci* **73:** 3852–3856. doi:10.1073/pnas.73.11.3852

Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. 2009. MENε/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci* **106:** 2525–2530. doi:10.1073/pnas.0807899106

Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ. 2017. Distinctive patterns of transcription and RNA processing for human lincRNAs. *Mol Cell* **65:** 25–38. doi:10.1016/j.molcel.2016.11.029

Sharma S, Lafontaine DLJ. 2015. 'View from a bridge': a new perspective on eukaryotic rRNA base modification. *Trends Biochem Sci* **40:** 560–575. doi:10.1016/j.tibs.2015.07.008

Shin Y, Brangwynne CP. 2017. Liquid phase condensation in cell physiology and disease. *Science* **357:** eaaf4382. doi:10.1126/science.aaf4382

Shukla CJ, McCorkindale AL, Gerhardinger C, Korthauer KD, Cabili MN, Shechner DM, Irizarry RA, Maass PG, Rinn JL. 2018. High-throughput identification of RNA nuclear enrichment sequences. *EMBO J* **37:** e98452. doi:10.15252/embj.201798452

Simon MD, Pinter SF, Fang R, Sarma K, Rutenberg-Schoenberg M, Bowman SK, Kesner BA, Maier VK, Kingston RE, Lee JT. 2013. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504:** 465–469. doi:10.1038/nature12719

Sloan KE, Leisegang MS, Doebele C, Ramirez AS, Simm S, Safferthal C, Kretschmer J, Schorge T, Markoutsa S, Haag S, et al. 2015. The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21. *Nucleic Acids Res* **43:** 553–564. doi:10.1093/nar/gku1291

Smola MJ, Rice GM, Busan S, Siegfried NA, Weeks KM. 2015. Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat Protoc* **10:** 1643–1669. doi:10.1038/nprot.2015.103

Souquere S, Beauclair G, Harper F, Fox A, Pierron G. 2010. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol Biol Cell* **21:** 4020–4027. doi:10.1091/mbc.e10-08-0690

Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung LH, Bindereif A. 2015. Exon circularization requires canonical splice signals. *Cell Rep* **10:** 103–111. doi:10.1016/j.celrep.2014.12.002

Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. 2009. MEN ε/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19:** 347–359. doi:10.1101/gr.087775.108

Szczurek A, Xing J, Birk UJ, Cremer C. 2016. Single molecule localization microscopy of mammalian cell nuclei on the nanoscale. *Front Genet* **7:** 114. doi:10.3389/fgene.2016.00114

Tafforeau L, Zorbas C, Langhendries JL, Mullineux ST, Stamatopoulou V, Mullier R, Wacheul L, Lafontaine DL. 2013. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of pre-rRNA processing factors. *Mol Cell* **51:** 539–551. doi:10.1016/j.molcel.2013.08.011

Talhouarne GJS, Gall JG. 2014. Lariat intronic RNAs in the cytoplasm of *Xenopus tropicalis* oocytes. *RNA* **20:** 1476–1487. doi:10.1261/rna.045781.114

Tessarz P, Santos-Rosa H, Robson SC, Sylvestersen KB, Nelson CJ, Nielsen ML, Kouzarides T. 2014. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505:** 564–568. doi:10.1038/nature12819

Thiry M, Lafontaine DL. 2005. Birth of a nucleolus: the evolution of nucleolar compartments. *Trends Cell Biol* **15:** 194–199. doi:10.1016/j.tcb.2005.02.007

Thiry M, Cheutin T, O'Donohue MF, Kaplan H, Ploton D. 2000. Dynamics and three-dimensional localization of ribosomal RNA within the nucleolus. *RNA* **6:** 1750–1761. doi:10.1017/S1355838200001564

Tichon A, Gil N, Lubelsky Y, Havkin Solomon T, Lemze D, Itzkovitz S, Stern-Ginossar N, Ulitsky I. 2016. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat Commun* **7:** 12209. doi:10.1038/ncomms12209

Tichon A, Perry RB, Stojic L, Ulitsky I. 2018. SAM68 is required for regulation of Pumilio by the NORAD long noncoding RNA. *Genes Dev* **32:** 70–78. doi:10.1101/gad.309138.117

Tollervey D, Lehtonen H, Carmo-Fonseca M, Hurt EC. 1991. The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J* **10:** 573–583. doi:10.1002/j.1460-2075.1991.tb07984.x

Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC. 1993. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* **72:** 443–457. doi:10.1016/0092-8674(93)90120-F

Tomita S, Abdalla MO, Fujiwara S, Yamamoto T, Iwase H, Nakao M, Saitoh N. 2017. Roles of long noncoding RNAs in chromosome domains. *Wiley Interdiscip Rev RNA* **8:** e1384. doi:10.1002/wrna.1384.

Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, et al. 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39:** 925–938. doi:10.1016/j.molcel.2010.08.011

Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA. 2012. Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Rep* **2:** 26–32. doi:10.1016/j.celrep.2012.05.020

Ulitsky I. 2016. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat Rev Genet* **17:** 601–614. doi:10.1038/nrg.2016.85

Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigo R, Johnson R. 2018. Towards a complete map of the human long noncoding RNA transcriptome. *Nat Rev Genet* **19:** 535–548. doi:10.1038/s41576-018-0017-y

Wang Y, Wang ZF. 2015. Efficient backsplicing produces translatable circular mRNAs. *RNA* **21:** 172–179. doi:10.1261/rna.048272.114

Wang PL, Bao Y, Yee MC, Barrett SP, Hogan GJ, Olsen MN, Dinneny JR, Brown PO, Salzman J. 2014. Circular RNA is expressed across the eukaryotic tree of life. *PLoS One* **9:** e90859. doi:10.1371/journal.pone.0090859

Wang Y, Hu SB, Wang MR, Yao RW, Wu D, Yang L, Chen LL. 2018. Genome-wide screening of *NEAT1* regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat Cell Biol* **20:** 1145–1158. doi:10.1038/s41556-018-0204-2

Weinstein LB, Steitz JA. 1999. Guided tours: from precursor snoRNA to functional snoRNP. *Curr Opin Cell Biol* **11:** 378–384. doi:10.1016/S0955-0674(99)80053-2

Wesselhoeft RA, Kowalski PS, Parker-Hale FC, Huang YX, Bisaria N, Anderson DG. 2019. RNA circularization diminishes immunogenicity and can extend translation duration in vivo. *Mol Cell* **74:** 508–520. doi:10.1016/j.molcel.2019.02.015

Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR, Lai EC. 2014. Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep* **9:** 1966–1980. doi:10.1016/j.celrep.2014.10.062

Wilusz JE, Freier SM, Spector DL. 2008. 3′ end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135:** 919–932. doi:10.1016/j.cell.2008.10.012

Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* **26:** 2392–2407. doi:10.1101/gad.204438.112

Wu H, Yin QF, Luo Z, Yao RW, Zheng CC, Zhang J, Xiang JF, Yang L, Chen LL. 2016. Unusual processing generates SPA LncRNAs that sequester multiple RNA binding proteins. *Mol Cell* **64:** 534–548. doi:10.1016/j.molcel.2016.10.007

Wu H, Yang L, Chen LL. 2017. The diversity of long noncoding RNAs and their generation. *Trends Genet* **33:** 540–552. doi:10.1016/j.tig.2017.05.004

Xia P, Wang S, Ye BQ, Du Y, Li C, Xiong Z, Qu Y, Fan ZS. 2018. A circular RNA protects dormant hematopoietic stem cells from DNA sensor cGAS-mediated exhaustion. *Immunity* **48:** 688–701. doi:10.1016/j.immuni.2018.03.016

Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, et al. 2014. Human colorectal cancer-specific *CCAT1-L* lncRNA regulates long-range chromatin interactions at the *MYC* locus. *Cell Res* **24:** 513–531. doi:10.1038/cr.2014.35

Xing YH, Chen LL. 2018. Processing and roles of snoRNA-ended long noncoding RNAs. *Crit Rev Biochem Mol Biol* **53:** 596–606. doi:10.1080/10409238.2018.1508411

Xing YH, Yao RW, Zhang Y, Guo CJ, Jiang S, Xu G, Dong R, Yang L, Chen LL. 2017. *SLERT* regulates DDX21 rings associated with Pol I transcription. *Cell* **169:** 664–678.e616. doi:10.1016/j.cell.2017.04.011

Yamazaki T, Souquere S, Chujo T, Kobelke S, Chong YS, Fox AH, Bond CS, Nakagawa S, Pierron G, Hirose T. 2018. Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol Cell* **70:** 1038–1053.e1037. doi:10.1016/j.molcel.2018.05.019

Yang L, Duff MO, Graveley BR, Carmichael GG, Chen LL. 2011. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol* **12:** R16. doi:10.1186/gb-2011-12-2-r16

Yang Y, Fan XJ, Mao MW, Song XW, Wu P, Zhang Y, Jin YF, Yang Y, Chen LL, Wang Y, et al. 2017. Extensive translation of circular RNAs driven by N^6-methyladenosine. *Cell Res* **27:** 626–641. doi:10.1038/cr.2017.31

Yao RW, Wang Y, Chen LL. 2019a. Cellular functions of long noncoding RNAs. *Nat Cell Biol* **21:** 542–551. doi:10.1038/s41556-019-0311-8

Yao RW, Xu G, Wang Y, Shan L, Luan PF, Wang Y, Wu M, Yang LZ, Xing YH, Yang L, et al. 2019b. Nascent Pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. *Mol Cell* **76:** 767–783. doi:10.1016/j.molcel.2019.08.014

Yin QF, Yang L, Zhang Y, Xiang JF, Wu YW, Carmichael GG, Chen LL. 2012. Long noncoding RNAs with snoRNA ends. *Mol Cell* **48:** 219–230. doi:10.1016/j.molcel.2012.07.033

Ying Y, Wang XJ, Vuong CK, Lin CH, Damianov A, Black DL. 2017. Splicing activation by Rbfox requires self-aggregation

through its tyrosine-rich domain. *Cell* **170:** 312–323.e310. doi:10.1016/j.cell.2017.06.022

You XT, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang MT, Glock C, Quedenau C, et al. 2015. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat Neurosci* **18:** 603–610. doi:10.1038/nn.3975

Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, et al. 2012. The lncRNA *Malat1* is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep* **2:** 111–123. doi:10.1016/j.celrep.2012.06.003

Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu SS, Yang L, Chen LL. 2013. Circular intronic long noncoding RNAs. *Mol Cell* **51:** 792–806. doi:10.1016/j.molcel.2013.08.017

Zhang XO, Wang HB, Zhang Y, Lu XH, Chen LL, Yang L. 2014a. Complementary sequence-mediated exon circularization. *Cell* **159:** 134–147. doi:10.1016/j.cell.2014.09.001

Zhang XO, Yin QF, Wang HB, Zhang Y, Chen T, Zheng P, Lu X, Chen LL, Yang L. 2014b. Species-specific alternative splicing leads to unique expression of *sno-lncRNAs*. *BMC Genomics* **15:** 287. doi:10.1186/1471-2164-15-287

Zhang XO, Dong R, Zhang Y, Zhang JL, Luo Z, Zhang J, Chen LL, Yang L. 2016a. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res* **26:** 1277–1287. doi:10.1101/gr.202895.115

Zhang Y, Xue W, Li X, Zhang J, Chen SY, Zhang JL, Yang L, Chen LL. 2016b. The biogenesis of nascent circular RNAs. *Cell Rep* **15:** 611–624. doi:10.1016/j.celrep.2016.03.058

Zhang SY, Clark NE, Freije CA, Pauwels E, Taggart AJ, Okada S, Mandel H, Garcia P, Ciancanelli MJ, Biran A, et al. 2018. Inborn errors of RNA lariat metabolism in humans with brainstem viral infection. *Cell* **172:** 952–965. doi:10.1016/j.cell.2018.02.019

Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, Li Z, Bu D, Sun N, Zhang MQ, et al. 2016. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res* **44:** D203–D208. doi:10.1093/nar/gkv1252

Zheng XF, Bevilacqua PC. 2004. Activation of the protein kinase PKR by short double-stranded RNAs with single-stranded tails. *RNA* **10:** 1934–1945. doi:10.1261/rna.7150804