

The nuclear functions of long noncoding RNAs come into focus

Zhenxing Song^{a,b,1}, Jiamei Lin^{a,b,1}, Zhengguo Li^{a,b}, Chuan Huang^{a,b,*}

^a School of Life Sciences, Chongqing University, Chongqing, 401331, China

^b Center of Plant Functional Genomics, Institute of Advanced Interdisciplinary Studies, Chongqing University, Chongqing, 401331, China

ARTICLE INFO

Keywords:

MALAT1

NEAT1

XIST

Circular RNA

RNA-Binding protein

Nucleus

ABSTRACT

Long noncoding RNAs (lncRNAs), defined as untranslated and tightly-regulated transcripts with a length exceeding 200 nt, are common outputs of the eukaryotic genome. It is becoming increasingly apparent that many lncRNAs likely serve as important regulators in a variety of biological processes. In particular, some of them accumulate in the nucleus and function in diverse nuclear events, including chromatin remodeling, transcriptional regulation, RNA processing, DNA damage repair, etc. Here, we unite recent progresses on the functions of nuclear lncRNAs and provide insights into the future research directions of this field.

1. Introduction

It was long assumed that only a small proportion of the eukaryotic genome is actively transcribed to generate templates for protein translation. However, RNA deep sequencing of eukaryotic transcriptomes has revealed that transcription is pervasively existing across eukaryotic genomes and a diverse group of RNA classes with little or no protein-coding capacity are produced [1,2]. Among these transcripts, long noncoding RNAs (lncRNAs) are ranging from 200 nt to 100 kb in length. Although most lncRNAs (e.g. X-inactive specific transcript, aka *XIST*) are generated with splicing signals and/or modifications similar to that of protein-coding genes as described by a comprehensive analysis [3], a number of lncRNAs are processed in unexpected ways. For example, the 3' end of *MALAT1* (Metastasis associated lung adenocarcinoma transcript 1) and *NEAT1* (Nuclear enriched abundant transcript 1) are processed by the tRNA biogenesis machinery, generating a stable triple helix structure instead of a poly(A) tail [4–6]. Circular RNAs are generated via back-splicing, a reaction by which a splicing donor is joined to an upstream acceptor [7–9]. Interestingly, some of lncRNAs are highly conserved at the sequence level. The expression levels of certain lncRNAs are higher than most protein-coding genes in a cell- and/or tissue-specific manner. Importantly, the dysregulation of many lncRNAs are associated with various human diseases, such as cancers and brain disorders. For example, *H19* is one of the most abundant lncRNAs during fetal development in mammals [10,11]. *H19* is robustly induced during embryogenesis at a level comparable to the

house-keeping gene β -actin, but is downregulated significantly after birth. Nevertheless, it remains highly expressed in adult skeletal muscle and heart [10,11]. Loss of *H19* induces precocious muscle differentiation and defective embryogenesis [10,11].

Similar to mRNAs, lots of lncRNAs display a post-transcriptional role in the cytoplasm as well summarized by Rashid et al. [12]. For example, cytoplasmic *lincRNA-p21* has an inhibitory effect on the translation activity of *CTNNB1* and *JUNB* by sterically blocking their transcripts [13]. In addition, circular RNA *ND1* and *ND5* can promote the mitochondrial entry of nuclear-encoded proteins by interacting with the mitochondrial importation complex [14]. But it is important to note that there are also many lncRNAs accumulating and functioning in the nucleus, which has gained a broad interest in recent years. These nuclear lncRNAs exert critical roles in chromatin remodeling, transcriptional regulation, RNA processing and modification, etc. In this review, we will discuss the functions and underlying mechanisms of several well-studied nuclear lncRNAs and provide insights into the future research directions of nuclear lncRNAs in the field.

2. Nuclear speckle-associated lncRNA *MALAT1* (metastasis associated lung adenocarcinoma transcript 1)

The *MALAT1* locus is located on human chromosome 11q13 and is able to generate a long non-coding and nuclear-retained transcript with a length of ~8 kb [15–17]. Despite being transcribed by RNA polymerase II (Pol II) [18,19], mature *MALAT1* transcript does not contain a

* Corresponding author. School of Life Sciences, Chongqing University, Chongqing, 401331, China.

E-mail address: chuanhuang@cqu.edu.cn (C. Huang).

¹ These authors contribute equally.

<https://doi.org/10.1016/j.ncrna.2021.03.002>

Received 4 February 2021; Received in revised form 18 March 2021; Accepted 22 March 2021

Available online 27 March 2021

2468-0540/© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC

BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

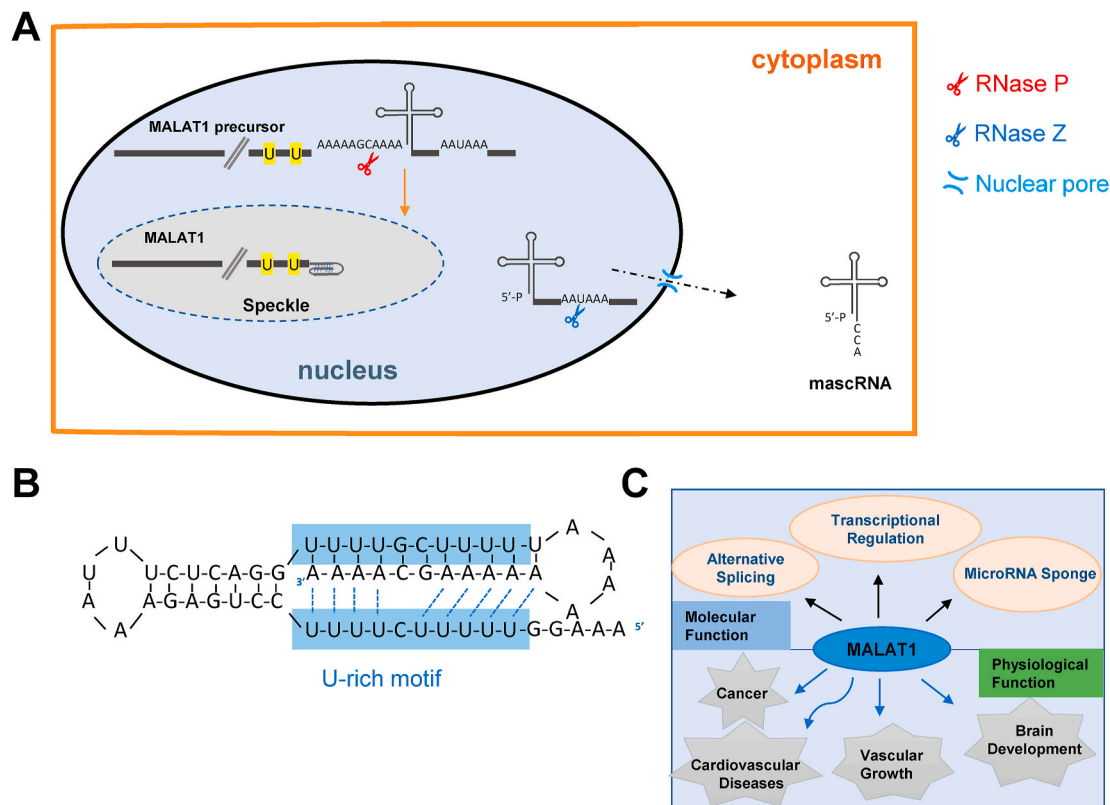


Fig. 1. The biogenesis, structure, and functions of MALAT1. (A) The MALAT1 precursor is first cleaved by RNase P immediately downstream of the poly(A)-rich tract, generating the 3' end of MALAT1 and the 5' end of mascRNA. The 3' end of mascRNA is then cleaved by RNase Z and added with CCA. MALAT1 accumulates in nuclear speckles, while mascRNA is exported to the cytoplasm. (B) The A- and U-rich motifs at the 3' end of MALAT1 form a triple helix via base-pairing, protecting the 3' end of MALAT1 from degradation. Figure is adapted from Ref. [5]. (C) MALAT1 participates in diverse molecular events and is associated with many diseases.

canonical poly(A) tail, since the 3' end processing of MALAT1 is completely different from the classical mRNA 3' terminus formation [4, 20,21]. The 3' end of MALAT1 precursor lacks canonical cleavage/polyadenylation signals, but instead consists of a short and highly conserved poly(A)-rich tract [4]. During generation of mature MALAT1, the precursor is first cleaved endonucleolytically by RNase P immediately downstream of the poly(A)-rich tract to simultaneously produce the 3' end of MALAT1 and the 5' end of a tRNA-like small RNA (MALAT1-associated small cytoplasmic RNA, aka mascRNA) [4]. The 3' end of mascRNA is then cleaved by RNase Z, followed by CCA addition (Fig. 1A) [4]. Curiously, MALAT1 accumulates in the nucleus, whereas mascRNA mainly translocates to the cytoplasm (Fig. 1A) [4]. Although MALAT1 lacks a poly(A) tail, it is highly abundant in human and mouse tissues [16,22,23] and expressed at a level higher than a number of house-keeping genes [24], indicating that the 3' end of MALAT1 is able to resist 3' degradation by 3'-5' exonucleases. Indeed, a following study found that the highly conserved A- and U-rich motifs at the 3' end of MALAT1 can form a triple helix via base-pairing, and disrupting the triple helix triggers the degradation of MALAT1 (Fig. 1B) [5,6,25], demonstrating that this extensive secondary structure is able to protect mature MALAT1 from 3'-5' exonucleolytic decay.

Different from mascRNA or many other ncRNAs, MALAT1 specifically localizes in nuclear speckles (also called as interchromatin granule cluster) [23,26,27], nuclear domains that are composed of multiple factors involved in transcription activation or repression, pre-mRNA splicing, mRNA modification, and mRNA nuclear export [28–30]. Three nuclear speckle proteins (RNPS1, IBP160, and SRm160) and two distinct motifs on MALAT1 contribute to the specific localization [27]. Such localization suggests that MALAT1 may play important roles in some molecular processes in the nucleus. In fact, MALAT1 is not a structural RNA required for the maintenance of nuclear speckle integrity

[31,32], but functions in chromosome segregation during mitosis [32]. Lack of MALAT1 results in aberrant mitosis and the fragmentation of cell nuclei [32]. Intriguingly, Tripathi and colleagues found that MALAT1 interacts with the serine/arginine (SR) splicing factors, controls the distribution and phosphorylation status of SR proteins in nuclear speckles and, in turn, modulates alternative splicing (AS) events of a set of SR-regulated pre-mRNAs in HeLa cells (Fig. 1C) [32]. It is thus likely that the cycling of at least some of SR proteins between speckle domains and sites of transcription is regulated by MALAT1. However, there are several contradictory observations that MALAT1 had almost no effect on the localization of nuclear speckle components or phosphorylation status of SR proteins in knockout mice [24,33,34]. Besides, no obvious connection between MALAT1 and AS regulation was found [24,33]. These inconsistent results might be explained by the possibility that MALAT1 has evolved a function in human cells that is not easily detected in mice or the function of MALAT1 is simply cell/tissue-specific, since MALAT1 expression varies significantly among tissues and cells [16,33].

MALAT1 has also been reported to participate in modulation of transcription (Fig. 1C). Indeed, a comprehensive analysis of genomic binding sites of MALAT1 implied that it exhibits localization to many actively transcribed chromatin sites [35]. The human *LTBP3* is a gene downstream of the MALAT1 locus and has an opposite transcription orientation to that of MALAT1. MALAT1 activates the transcription of *LTBP3* by promoting the recruitment of the key transcription factor Sp1 to the *LTBP3* promoter in mesenchymal stem cells [36]. A large reduction in expression of *NEAT1*, an lncRNA locus upstream of MALAT1, was observed in MEFs obtained from MALAT1 knockout mice [27]. In addition to the *cis* effect on neighboring genes, the localization of the polycomb 2 protein (Pc2), a key histone modifier, is controlled by MALAT1, which in turn influences expression of many genes [37]. Yang and colleagues demonstrated that the association between MALAT1 and

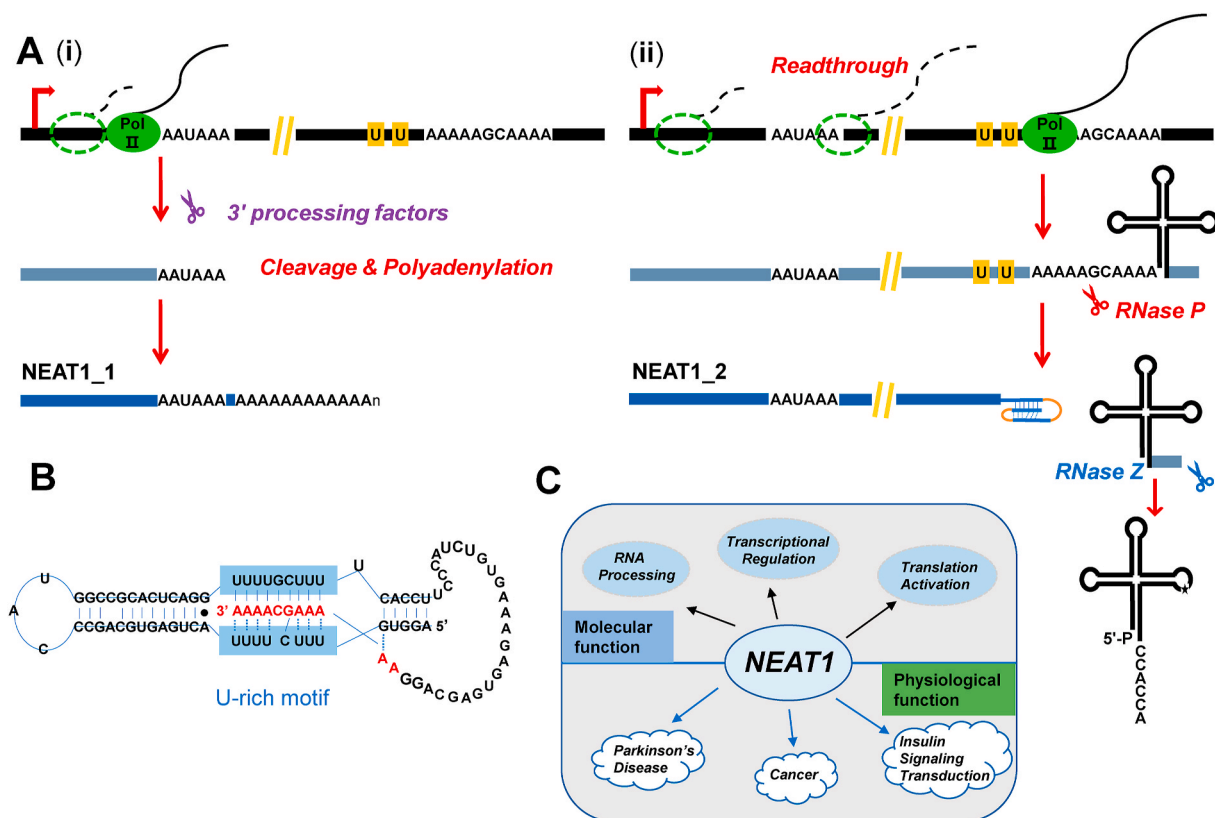


Fig. 2. The biogenesis, structure, and functions of NEAT1. (A) NEAT1_1 (the short form of NEAT1) is stabilized by a canonical poly(A) tail following cleavage/polyadenylation. However, NEAT1_2 (the long form of NEAT1) undergoes steps similar to that of MALAT1 biogenesis. (B) The A- and U-rich motifs, present upstream of the RNase P cleavage site, form a triple helix via base-pairing, protecting the 3' end of NEAT1_2 from degradation. Figure is adapted from Ref. [6]. (C) NEAT1 is involved in diverse molecular events, such as RNA processing and transcriptional regulation. The dysregulation of NEAT1 contributes to many diseases.

unmethylated Pc2 promotes E2F1 SUMOylation, resulting in a biochemical switch of the Pc2 chromodomain from interacting with repressive histone marks to preferring active histone marks. Subsequently, a number of growth-control genes are activated by the event [37].

MALAT1 can also act as a microRNA sponge in the nucleus in an AGO2-dependent manner (Fig. 1C) [38–41]. For example, there are two targeting sites for *mircoRNA-9* on MALAT1 transcript. The MALAT1 level was significantly reduced in AGO2 immunoprecipitates purified from cells treated with anti-*mircoRNA-9* as compared to the control sample [38]. In another case, MALAT1 can modulate *c-Myc* expression by sponging *mircoRNA-34a* in melanoma cells [40].

MALAT1 transcript was first discovered in a screening for genes whose expression have a potential relationship to metastasis in early-stage non-small cell lung cancer [16]. Afterwards, MALAT1 was found to be evolutionally conserved from mouse to human and widely expressed in numerous cancer tissues [42–45], implying that the dysregulation of MALAT1 plays crucial roles in cancers and MALAT1 can be used as a valid cancer biomarker as well as a therapeutic target for clinical application (Fig. 1C). For instance, MALAT1 has been found as an important regulator of lung cancer metastasis probably due to its effect on metastasis-associated gene expression [46]. Interestingly, the MALAT1 ASO (antisense oligonucleotide) treated mice was able to resist to EBC-1 tumor metastasis to the lung, indicating a potential application of ASOs against MALAT1 in preventing lung cancer from spreading [46]. In colorectal cancer (CRC), the binding of MALAT1 to the tumor suppressor PSF (PTB-associated splicing factor) dissociates PTB (polypyrimidine-tract-binding protein) from PSF. The released PTB promotes CRC cell proliferation and tumor growth [47]. Two recent studies indicated that the secondary structure of MALAT1 varies in different cancers [48,49]. By dimethyl sulfate-sequencing and RNA structure

analysis, MALAT1 is likely to be more unstructured in chronic myeloid leukemia, but undergoes 18 novel structural rearrangements in cervical cancer [49]. This provides new insights into how RNA structure affects its regulatory role under various conditions. Besides cancers, the dysregulation of MALAT1 is also involved in many other physiological processes and pathological diseases, including brain development [26, 50], vascular growth [51], cardiovascular diseases [52,53], neurological disorders [54–57], just to name a few. Unfortunately, most of the precise mechanisms underlying the MALAT1-associated diseases still remain unclear. Therefore, further studies are needed to clarify these questions.

3. Nuclear paraspeckle-associated lncRNA NEAT1 (nuclear enriched abundant transcript 1)

NEAT1 is originally identified as a large and infrequently spliced transcript in a genome-wide screening to search for nuclear ncRNAs that might modulate gene expression [23]. The NEAT1 locus is adjacent to the MALAT1 locus on human chromosome 11q13 and generates two overlapping isoforms [58,59]. NEAT1_1, aka *MEN α* , is the short isoform with a length of ~3700 nucleotides [60]. The 3' end of NEAT1_1 undergoes the canonical cleavage/polyadenylation steps and, in turn, terminates in a poly(A) tail (Fig. 2A). On the contrary, the long version of NEAT1 (NEAT1_2, aka *MEN β*) is over 23 kb and stabilized by a triple helix structure following RNase P cleavage at its 3' end (Fig. 2A and B) [5,6,60], which is similar to MALAT1. Although none of the two isoforms has an open reading frame or a capacity to yield a peptide, they are highly and ubiquitously expressed in a wide range of tissues, such as ovary, prostate, colon, and pancreas [23]. As a type of architectural lncRNA, NEAT1 is essential for organization and integrity of nuclear paraspeckles [61–64], irregularly shaped subnuclear bodies closed to the NEAT1 transcription site [31]. NEAT1 translocates immediately to

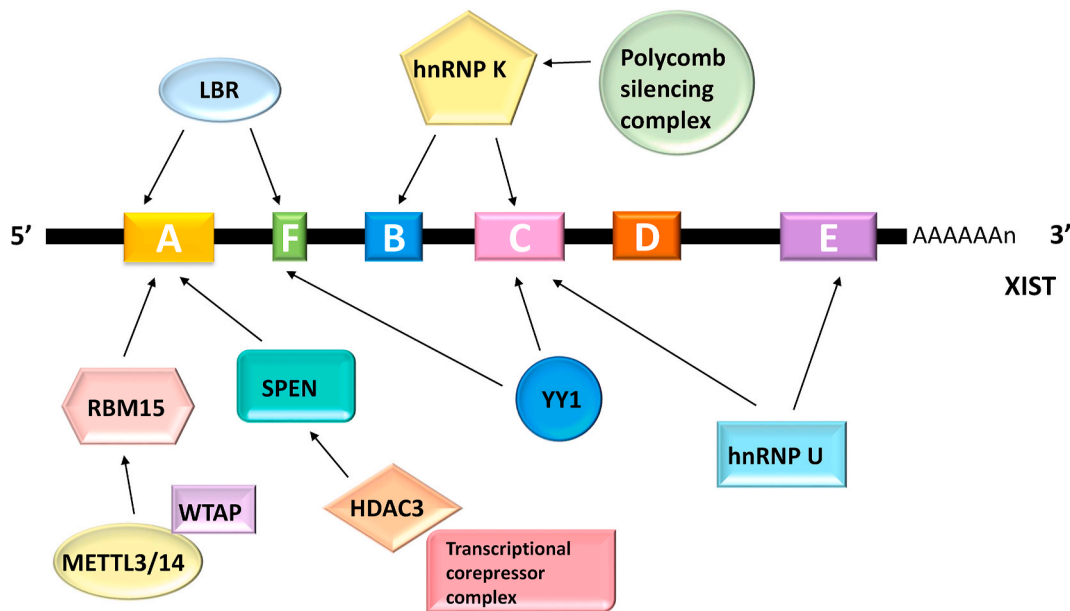


Fig. 3. Schematic representation of XIST and its RBPs. As a multi-tasking RNA molecule, XIST functions in a variety of molecular processes through different RBPs. The tandem repeats (denoted in boxes) contribute to the recruitment of different RBPs.

nuclear paraspeckles once it is transcribed by RNA Pol II [65]. During daughter nuclei reformation, *NEAT1* transcription happens prior to the accumulation of the paraspeckle-associated proteins PSP1/p54 in paraspeckles [31]. *NEAT1* depletion leads to loss of paraspeckles, whereas overexpression of *NEAT1* increases paraspeckle number [31,60,65,66], implying that the paraspeckle-associated proteins might be multifunctional when endogenous *NEAT1* expression is altered. In fact, *NEAT1* has been found to bind specifically to PSP1/p54 [31]. Interestingly, PSP1 forms a heterodimer with p54 in a *NEAT1*-independent manner [66]. But PSP1 without the RNA recognition motifs no longer exhibits an interaction with *NEAT1* or paraspeckles [31]. Furthermore, *NEAT1* depletion has no significant effect on the protein level of PSP1 or p54, excluding the possibility that *NEAT1* regulates the degradation of some paraspeckle proteins [60]. Therefore, nuclear paraspeckle is built on *NEAT1* and the dysfunction of nuclear paraspeckle could be triggered by *NEAT1* depletion. For example, the nuclear export of transcripts containing paraspeckle-retention elements (*Alu* repeats in humans and *SINEs* in mouse) or hyper-edited A-to-I is enhanced by *NEAT1* depletion [66,67].

NEAT1 not only functions as an architectural component of paraspeckles, but also is important to some other organelles (Fig. 2C). For example, *NEAT1* and paraspeckles are crucial for mitochondrial homeostasis [68,69]. A recent study demonstrated that loss of mitochondrial proteins or mitochondrial stress results in abnormal expression of *NEAT1* as well as a remarkable increase of elongated nuclear paraspeckles [68]. Moreover, many mRNAs of nuclear-encoded mitochondrial proteins contain paraspeckle-retention elements at their 3' UTRs and is associated with *NEAT1* [68], indicating a strong regulatory communication between nucleus and mitochondria. Abnormal *NEAT1* expression leads to mitochondrial dysfunction, such as altered mitochondrial respiration and ATP production rate [68,69].

NEAT1 has strong regulatory effects on many molecular processes (Fig. 2C). Jiang and colleagues found that *NEAT1* and its binding partner PSP1/p54 heterodimer can enhance pri-miRNA processing [70]. *NEAT1* may provide a binding platform for various pri-miRNAs and microprocessor components that are attracted by the pri-miRNA-like hairpin structures of *NEAT1* [70]. Additionally, *NEAT1* can regulate the splicing of *PPAR γ 2* (peroxisome proliferator activated receptor) pre-mRNA temporally via the SR protein SRp40 mediated pathway during adipogenesis. In this case, *NEAT1* binds to SRp40 and potentiates its

phosphorylation status by Clk kinase [71]. Wang and colleagues found that *NEAT1* can activate viral gene expression by recruiting STAT3 (signal transducer and activator of transcription 3) to their promoters in response to herpes simplex virus-1 (HSV-1) infection [72]. Similarly, their following study showed that *NEAT1* recruits the acetyltransferase CBP/P300 complex to the promoters of the endocytosis-related genes, thereby altering H3K27 acetylation and H3K27 crotonylation in Alzheimer disease [73].

Except for RNA processing or transcriptional regulation, several studies have suggested that *NEAT1* has a profound connection with translation activation or protein stability control. For example, during RNA polymerase I inhibition, *NEAT1*-sequestered p54 and PSF are released from nuclear paraspeckles upon *NEAT1* depletion and relocated to *c-Myc* mRNAs for its translation activation [74]. Curiously, *NEAT1* depletion has almost no effect on the protein level of *c-Myc* under normal conditions [74], indicating that the *NEAT1*-mediated *c-Myc* translation only occurs in stressed conditions. In CD4⁺ T cells, *NEAT1* physically binds to STAT3 protein and keeps it from ubiquitination [75]. This was exemplified by the *NEAT1* knockdown experiment which led to a large reduction in ubiquitination labeled STAT3 protein [75]. Likewise, the half-life of another *NEAT1*-binding protein DDX5 is drastically increased upon overexpression of *NEAT1* in HT29 cells [76].

Similar to many other ncRNAs, *NEAT1* consists of potential targeting sites for various miRNAs [77–82]. Importantly, plenty of these miRNA-target genes are oncogenes and contribute to tumorigenesis. Disrupting the cross-talk between *NEAT1*, miRNAs, and miRNA-target genes may promote proliferation, migration, invasion of cancer cells. Thus *NEAT1* could be an effective target for the diagnosis and therapy of many cancer types. In prostate cancer, *NEAT1* modulates the expression of the oncogene *HMG2* by sponging *miR-98-5p* [81]. In gastric cancer, the binding of *miR-365a-3p* to *NEAT1* enhances *ABCC4* expression [80]. In esophageal Squamous Cell Carcinoma, *miR-129/CTBP2* axis is controlled by *NEAT1* [82]. Nevertheless, it is still elusive how and why cytoplasmic miRNAs are controlled by a nuclear lncRNA.

An increasing number of studies focusing on *NEAT1*'s function in physiology have proved that the intensive upregulation of *NEAT1* in several human diseases is associated with a drastic decrease in patient survival (Fig. 2C) [83–87]. In regard to Parkinson's disease (PD), a chronic and progressive brain disorder that affects movement, *NEAT1*

expression was found to be significantly upregulated in PD mice [83,84]. Interestingly, depletion of *NEAT1* could reverse apoptosis induced by PD inducers possibly by controlling the expression of some PD contributors, such as α -synuclein and PINK1 [83,84]. Regarding non-alcoholic fatty liver disease, loss of *NEAT1* inhibits the expression of acetyl-CoA carboxylase and fatty acid synthase through the mTOR/S6K1 signaling pathway, implying that *NEAT1* acts as a potent regulator in the biosynthesis of lipid and insulin signaling transduction [88]. With intensive researches of *NEAT1*-related diseases, we believe that a therapeutic application of *NEAT1* is coming soon.

4. Chromatin-associated lncRNA *XIST* (X-inactive specific transcript)

Less than 3% of the ancestral genes survive on the male Y chromosome, while X chromosome retains more than 97% [89,90]. Therefore, males and females have unbalanced genetic content in mammals. To minimize sex chromosome-linked gene expression imbalance between XX females and XY males, a randomly chosen X chromosome is transcriptionally silenced in early female embryogenesis. This dosage compensation process is called X chromosome inactivation (XCI) and is initiated by the lncRNA *XIST* [91–96]. It is important to note that XCI by no means equal to loss of X chromosome, since 3%–20% of the X-linked genes from the inactive X (Xi) chromosome are still activated and indispensable for female embryogenesis. This is exemplified by Turner syndrome, a genetic disorder disease only occurs in females with only a single X chromosome [97,98]. *XIST* was first described as a long non-coding transcripts generated exclusively from the Xi chromosome in the early 1990s [99–103]. Once generated, *XIST* spreads across the entire length of the Xi chromosome, resulting in gene repression chromosome-widely. Further sequence analyses revealed that *XIST* is a 15–17 kb spliced and polyadenylated RNA which contains blocks of local tandem repeat elements designated A-F (Fig. 3) [102,103]. Significant progresses have elucidated the functional roles of each individual repeat in establishment of X chromosome-linked gene silencing. Hence, we will focus our discussion on these functional elements as well as their RBPs (RNA-binding proteins) in this section.

Using a series of inducible *XIST* transgenes, Wutz and colleagues found that *XIST* mutations that lack A-repeat, consisting of 7–8 copies of a 24 nt GC-rich core sequence at the 5' end of *XIST*, abolish silencing activity and lead to a decrease in stability without influencing localization [104]. Furthermore, the transcription activity of *XIST* appears to be affected by deletion of the endogenous A-repeat region in DNA [105]. Besides affecting *XIST* itself, the role of A-repeat in chromosome silencing is also linked to its binding proteins. For example, A-repeat contributes to the interaction with RBM15 (RNA binding motif protein 15) which further targets the methyltransferase complex METTL3/14 to specific RNAs, including *XIST*, for addition of *N*⁶-methyladenosine (m⁶A modification) [106,107]. Importantly, loss of the m⁶A catalytic subunit or the m⁶A reader leads to a major silencing deficiency, indicating a role of m⁶A in *XIST*-mediated chromosome silencing [106]. Contradictorily, only a minor decrease in X silencing efficiency was observed in either RBM15 knockdown or knockout cells [106,108]. The phenomena might be explained by the finding of RBM15B, the RBM15 homolog, which shares redundant functions with RBM15 [106,109]. SPEN is another interactor of A-repeat [110]. Depletion of SPEN results in impaired silencing of the X chromosome-linked genes, such as *Fbxl17*, without disrupting *XIST* localization [107,108,110,111]. In addition, SPEN might initiate transcriptional repression by actively recruiting the transcriptional corepressor complex and the histone deacetylase HDAC3 [112–114]. Unlike RBM15 and SPEN, LBR (Lamin B receptor) has a widespread distribution across *XIST* [115,116]. The major binding site encompasses the entire A- and F-repeat as proved by LBR CLIP-seq analysis [115,116]. Losing the interaction between *XIST* and LBR leads to impaired chromosome silencing [108,116]. B- and C-repeat have been reported to be critical for *XIST* spreading as well as the

recruitment of hnRNP K (heterogeneous nuclear ribonucleoprotein K) and the polycomb silencing complexes to *XIST* [110,117,118]. Chu and colleagues found that defective chromosome silencing can be induced by depletion of hnRNP K [110]. With further IF-coFISH (combined immunofluorescence and RNA fluorescence *in situ* hybridization) experiments, they demonstrated that hnRNP K colocalizes with H2AK119ub and H3K27me3, the earliest chromatin modifications occurring in XCI [110]. Depletion of hnRNP K reduces the accumulation of H3K27me3 and H2AK119ub along Xi without influencing the global level of these modifications, suggesting Xi undergoes a specific modification. Indeed, hnRNP K depletion specifically dissociates *XIST* from the polycomb silencing complexes [110]. C-repeat, together with E-repeat, is also responsible for the binding of hnRNP U (heterogeneous nuclear ribonucleoprotein U) to *XIST* and, in turn, *XIST* localization to chromatin [108,110,119–121]. But whether correct *XIST* localization depends on a combination of additional factors is still elusive, since the effect of hnRNP U cannot be confirmed in certain cell types [122,123]. In fact, an interesting study demonstrated that YY1 protein, capable of binding both RNA and DNA, acts as an adaptor in docking C-repeat of the RNA onto the F-repeat region of the *XIST* locus [124]. By deleting D-repeat through CRISPR/Cas9 technique in HEK293T cells, a significantly reduction in *XIST* expression was observed, implying that D-repeat might play an important role in the stability or transcriptional activity of *XIST* [125]. Consequently, the silencing ability of *XIST* was abrogated as proved by the phenomenon that levels of some X chromosome-linked genes were upregulated [125].

3D structural reorganization is another critical pathway for establishment of Xi gene repression by *XIST*, together with SmcHD1 (Structural maintenance of chromosomes flexible hinge domain containing 1) and the polycomb factors [126–134]. In this mechanism, the X inactive genes are relocated from the periphery to the interior of chromatin during XCI, thereby being isolated from RNA Pol II and other general transcription factors [126].

XCI is not only a dosage compensation process happening during early female embryogenesis and related to X-linked diseases in females, but also a critical modifier of many other diseases including cancers and neurological disorders. For example, *XIST* could affect cancer development and progression by sponging various microRNAs, thereby influencing the expression of some oncogenes [135–137]. On the other hand, Yue and colleagues found that the level of *XIST* is significantly increased in AD (Alzheimer's disease) mice models, suggesting that *XIST* might be used as a diagnostic marker for AD [138].

5. Nuclear circular RNA

Protein-coding genes of the eukaryotic genome are typically split by one or multiple introns. During the production of a mature mRNA, introns must be removed and exons are joined together in a linear order, which is known as canonical splicing [139–141]. Alternatively, splicing events can also occur in a way that a splicing donor is joined to an upstream acceptor, thereby generating a covalently-closed RNA (circular RNA) [7,142–145]. Despite the first circular RNA being discovered in 1976 [146], only recently has it been appreciated that these covalently-closed molecules represent a common class of transcriptional outputs from numerous protein-coding genes in the eukaryotic system [147,148]. So far, many labs have been putting a lot of efforts into characterization of circular RNA biology, and thus obtained remarkable progresses as well reviewed in many papers [7–9,143,149–154]. In short, the biogenesis of circular RNA is initiated by base-pairing between the flanking intronic repeats as well as some RBPs, including QKI, Mbl, ADAR1, RBM20, FUS, etc [155–159]. The degradation of circular RNA is controlled by different exonucleases and the platform protein gawky (aka GW182) in various context [160–164]. Once generated, most circular RNAs are exported to the cytoplasm in a length-dependent manner with the assistance of the DExH/D-box helicase DDX39A and DDX39B [165,166]; however, a small amount of circular RNAs accumulate in the

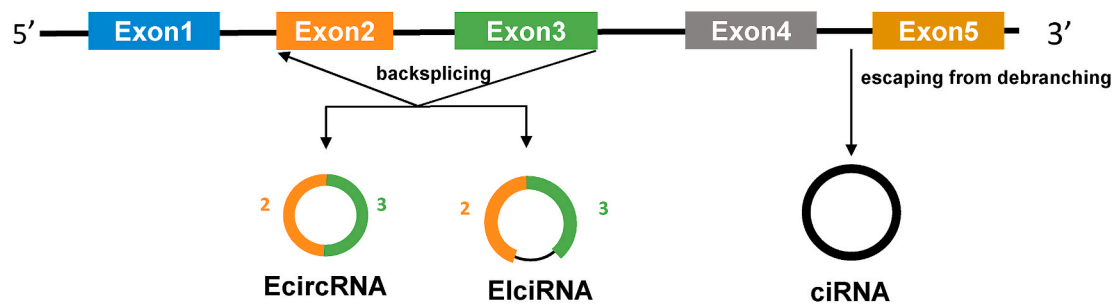


Fig. 4. The biogenesis of circular RNAs. Most circular RNAs are generated from protein-coding genes via back-splicing by which a splicing donor is joined to an upstream acceptor. In other cases, circular RNAs are derived from lariat intronic RNAs that escape from debranching.

Table 1
The nuclear functions of circular RNAs

circRNA	Type	Molecular Mechanism	Ref
circEIF3J	EiciRNA	Interacts with RNA Pol II and U1 snRNP to regulate transcription of its host gene	[167]
circPAIP2	EiciRNA	Interacts with RNA Pol II and U1 snRNP to regulate transcription of its host gene	[167]
circANKRD52	ciRNA	Interacts with RNA Pol II to regulate transcription of its host gene	[168]
circIns2/INS	ciRNA	Interacts with TDP-43 to regulate transcription	[169]
circFECD1	EciRNA	Induces promoter hypomethylation with the demethylase TET1	[170]
circ2082	EciRNA	Promotes DICER nuclear translocation via RBM3	[171]
cia-cGAS	EciRNA	Inhibits cGAS as a decoy	[172]
circSEP	EciRNA	Regulates alternative splicing of its host gene via R-loop	[173]
circCRM1	Intergenic circRNA	Maintains normal centromeric chromatin organization via R-loop	[174]
circSMARCA5	EciRNA	Downregulates its host gene expression via R-loop	[177]

nucleus. Here, we will focus our discussion on nuclear circular RNAs and their functions in this section.

In searching for lncRNAs involved in transcriptional regulation, Li and colleagues identified a special subclass of Pol II-binding circular RNA, also called EiciRNA (exon-intron circular RNA), whose intron(s) between circularized exons are not spliced out (Fig. 4) [167]. Similar to most EciRNAs (exonic circular RNAs), EiciRNAs are generated via back-splicing reactions in which the flanking *Alu* repeats facilitate their circularization [167]. Functionally, EiciRNAs have *cis* regulatory effects on their host gene transcription activity (Table 1), since depletion of EiciRNA leads to a significant reduction in the levels of their nascent mRNAs. Further mechanism analyses indicated that EiciRNA interacts with *U1* snRNP directly at the promoter of their host gene. Loss of *U1* snRNA or disrupting the *U1* snRNA-EiciRNA interaction by antisense morpholino (AMO) impairs the transcription-enhancing effect of EiciRNAs [167]. Notably, EiciRNAs not only localize to their sites of transcription, but also accumulate to other nuclear subdomains, suggesting other potential regulatory roles and mechanisms of EiciRNAs [167]. Another example of nuclear intron-containing circular RNA is ciRNA (intronic circular RNA), a type of lariat RNA which escapes from debranching (Fig. 4) [168]. Like EiciRNAs, ciRNAs (e.g. *circANKRD52*) are able to interact with the Pol II complex and modulate transcription efficiency of their host genes in *cis* (Table 1). Mechanistically, depletion of *circANKRD52* has almost no effect on Pol II engagement, suggesting that *circANKRD52* affects transcriptional rate but not transcription termination [168]. In another case, ciRNAs (*circIns2* in rat and *circINS* in human), derived from the *insulin* gene, regulate transcription of genes involved in calcium signaling and insulin exocytosis via the RBP TAR DNA-binding protein 43 kDa (TDP43) in the nucleus (Table 1) [169]. Low expression of *circIns2* or *circINS* leads to defective insulin secretion

of β -cell to and are associated with type 2 diabetes [169]. EciRNAs are also capable of regulating transcription in the nucleus as exemplified by *circFECD1*. *CircFECD1* can activate its host gene expression by inducing promoter hypomethylation with the demethylase TET1 (Table 1) [170].

Besides transcription, some nuclear circular RNAs can affect the function of their RBPs. For example, a recent study demonstrated that *circ2082* has an important role in nucleocytoplasmic distribution of the endoribonuclease DICER, a key component of pre-microRNA processor complex (Table 1) [171]. In this case, *circ2082* is recruited to DICER by RNA binding motif protein 3 (RBM3) in the nucleus of glioblastoma cell. The interaction is crucial for nuclear retention of DICER, since depletion of *circ2082* results in normalization of DICER localization to the cytoplasm [171]. The change in nucleocytoplasmic distribution of DICER further reestablishes the assortment of cellular microRNAs (microRNAome) [171]. Because the dysregulation of microRNAome is associated with cancers, *circ2082* could be used as a therapeutic target for cancer treatment. Circular RNA *cia-cGAS* act as a protein decoy for the DNA sensor cGAS and inactivates its synthase activity under homeostatic conditions, which in turn inhibits cGAS-mediated autoimmune responses (Table 1) [172].

Interestingly, certain nuclear circular RNAs are able to form an RNA:DNA hybrid (R-loop) with the genomic DNA. In *Arabidopsis*, nuclear *circSEP* derived from exon 6 of the *SEPALLATA3* gene was reported to affect the splicing of its host gene through an R-loop mechanism (Table 1). Formation of the R-loop leads to an increased level of a linear *SEP* transcript isoform which lacks exon 6, in turn driving floral homeotic phenotypes [173]. Importantly, *circSEP*:DNA R-loops are more abundant than their linear RNA:DNA counterparts, indicating the specific role of *circSEP* in R-loop formation [173]. In *Zea mays*, circular RNA:DNA R-loops are integral components of centromeric chromatin [174]. The centromeric retrotransposons derived circular RNAs were found to bind to the centromere through R-loops (Table 1). Disrupting the R-loop results in abnormal localization of the centromeric H3 variant (CENH3), in turn affecting centromeric chromatin organization [174]. SMARCA5 is a key component of DNA damage repair pathway and is required for chromatin remodeling [175,176]. In breast cancer cells, *circSMARCA5* can downregulate the expression of the *SMARCA5* gene via R-loop (Table 1). Therefore, *circSMARCA5* has a negative effect on DNA repair capacity [177].

6. Concluding remarks

Eukaryotic nucleus is a membrane-bound organelle contains most genetic information and controls gene expression. Many molecular processes, such as transcription, splicing, replication, histone modification, and DNA repair, occur in the nucleus and are regulated combinatorially by various protein factors. Only until recent 20 years, lncRNAs have been considered important regulators in the processes mentioned above [19,178–180]. Although some of lncRNAs (e.g. *XIST*) have general effects on nuclear events, the majority of them seem to act uniquely on certain genes, implying that the total numbers of functional

Table 2

The methods used to study RNA interactomes.

Method	Description	Experimental Purpose	Enhanced Version	Ref
CHIRP-seq	A method to pull down and map the RNA-associated chromatin using biotinylated antisense oligos	Identification of RNA-DNA interaction	CHART-seq, CHIRT-seq, hiCHIRP-seq	[181–185]
R-CHIP-seq	A method to pull down and map R-loops using catalytic-dead RNase H	Identification of R-loop		[186]
DRIP-seq	A method to pull down and map R-loops using S9.6 antibody	Identification of R-loop	bisDRIP-seq, DRIPc-seq	[187–189]
CLIP-seq	A method to pull down and map the protein-binding RNA by immunoprecipitation	Identification of RNA-protein interaction	PAR-CLIP-seq, iCLIP-seq, hiCLIP-seq	[190–193]
SHAPE-seq	A method to predict RNA secondary structures using chemical modifiers	Prediction of RNA-RNA interaction	icSHAPE-seq, DMS-seq, SHAPE-seq2.0	[194–197]
Frag-seq	A method to predict RNA secondary structures using endonuclease P1	Prediction of RNA-RNA interaction		[198]

nuclear lncRNAs (annotated or unannotated) may be underestimated. Thus defining the novel nuclear roles of lncRNAs will help us understand the complex of nuclear regulation.

Having established an association between a lncRNA and a nuclear event, we could reason a possible mechanism based on the phenotype. Informed by the studies discussed in above sections, we will summarize potential mechanisms for the functions of newly-discovered nuclear lncRNAs. First, lncRNAs might indirectly affect certain events via controlling the expression or subcellular localization of the key protein factors. With the development of high-throughput RNA-seq technique and single molecule tracking system, this possibility could be easily examined. Second, lncRNAs might act as a molecular platform, bringing diverse proteins together into an RNP complex, or a decoy by sponging RNAs or even proteins and preventing them from associating with their targets. In these two cases, deletion analyses can be performed. If the hypothesis is right, then deletion of a protein binding motif in lncRNAs or a RNA binding motif in proteins leads to loss of function. In addition, a variety of high-throughput sequencing methods were developed to identify the interactomes of nuclear lncRNAs *in vivo* (Table 2). These tools could help us map RNA-DNA [181–189], RNA-protein [190–193], and RNA-RNA interactions [194–198]. Third, lncRNAs might be an organizer of nuclear architecture. Knockdown or knockout of these lncRNAs results in nuclear subdomain breakdown, thereby affecting unique reactions. Fourth, the act of transcription through a lncRNA locus seems to be regulatory, rather than the lncRNA product as discussed by Long et al. and Bassett et al. [178,199]. To answer the question, the lncRNA region in genomic DNA can be replaced by a random sequence without changing its promoter. Combinatorially, the lncRNA transcript can be knockdown by siRNA, ASO, or CRISPR-Cas13a system or functionally blocked by AMO.

In conclusion, the last 20 years have revealed that lncRNAs are not simply rare transcriptional noise, but pervasively exist in the eukaryotic system. Although numerous surprises have been disclosed, they only represent a glimpse of the tip of lncRNA biology. More efforts are required to provide insights into how these special transcripts are controlled and function especially in the nucleus. The era for studies in lncRNA biology is just coming.

Declaration of competing interest

No potential conflict of interest was reported by the authors.

Acknowledgments

This work was sponsored by the National Natural Science Foundation of China (32070633), the Natural Science Foundation of Chongqing, China (cstc2019jcyjmsxmX0085), the Innovation Support Program for Overseas Returned Scholars of Chongqing, China (cx2019142), the Fundamental Research Funds for the Central Universities of China (2020CDJQY-A076), and the 100 Talent Program of Chongqing University (0304001104433). We thank the members of the Shan lab

(University of Science and Technology of China) and the Wilusz lab (University of Pennsylvania) for discussions.

References

- [1] P. Kapranov, A.T. Willingham, T.R. Gingeras, Genome-wide transcription and the implications for genomic organization, *Nat. Rev. Genet.* 8 (6) (2007) 413–423.
- [2] M.B. Clark, et al., The reality of pervasive transcription, *PLoS Biol.* 9 (7) (2011), e1000625 discussion e1001102.
- [3] T. Derrien, et al., The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, *Genome Res.* 22 (9) (2012) 1775–1789.
- [4] J.E. Wilusz, S.M. Freier, D.L. Spector, 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA, *Cell* 135 (5) (2008) 919–932.
- [5] J.E. Wilusz, et al., A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails, *Genes Dev.* 26 (21) (2012) 2392–2407.
- [6] J.A. Brown, et al., Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs, *Proc. Natl. Acad. Sci. U. S. A.* 109 (47) (2012) 19202–19207.
- [7] L. Chen, et al., Circular RNAs in eukaryotic cells, *Curr. Genom.* 16 (5) (2015) 312–318.
- [8] J.E. Wilusz, A 360 degrees view of circular RNAs: from biogenesis to functions, *Wiley Interdiscip. Rev. RNA* 9 (4) (2018) e1478.
- [9] M.S. Xiao, Y. Ai, J.E. Wilusz, Biogenesis and functions of circular RNAs come into focus, *Trends Cell Biol.* 30 (3) (2020) 226–240.
- [10] A. Gabory, et al., The H19 gene: regulation and function of a non-coding RNA, *Cytogenet. Genome Res.* 113 (1–4) (2006) 188–193.
- [11] A. Gabory, H. Jammes, L. Dandolo, The H19 locus: role of an imprinted non-coding RNA in growth and development, *Bioessays* 32 (6) (2010) 473–480.
- [12] F. Rashid, A. Shah, G. Shan, Long non-coding RNAs in the cytoplasm, *Dev. Reprod. Biol.* 14 (2) (2016) 73–80.
- [13] J.H. Yoon, et al., LincRNA-p21 suppresses target mRNA translation, *Mol. Cell* 47 (4) (2012) 648–655.
- [14] X. Liu, et al., Identification of meccRNAs and their roles in the mitochondrial entry of proteins, *Sci. China Life Sci.* 63 (10) (2020) 1429–1449.
- [15] R. Lin, et al., A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas, *Oncogene* 26 (6) (2007) 851–858.
- [16] P. Ji, et al., MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer, *Oncogene* 22 (39) (2003) 8031–8041.
- [17] X. Zhang, M.H. Hamblin, K.J. Yin, The long noncoding RNA Malat1: its physiological and pathophysiological functions, *RNA Biol.* 14 (12) (2017) 1705–1714.
- [18] S. Djebali, et al., Landscape of transcription in human cells, *Nature* 489 (7414) (2012) 101–108.
- [19] B. Yu, G. Shan, Functions of long noncoding RNAs in the nucleus, *Nucleus* 7 (2) (2016) 155–166.
- [20] J.E. Wilusz, D.L. Spector, An unexpected ending: noncanonical 3' end processing mechanisms, *RNA* 16 (2) (2010) 259–266.
- [21] J.E. Wilusz, Long noncoding RNAs: Re-writing dogmas of RNA processing and stability, *Biochim. Biophys. Acta* 1859 (1) (2016) 128–138.
- [22] K. Tano, et al., MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes, *FEBS Lett.* 584 (22) (2010) 4575–4580.
- [23] J.N. Hutchinson, et al., A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains, *BMC Genom.* 8 (2007) 39.
- [24] B. Zhang, et al., The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult, *Cell Rep.* 2 (1) (2012) 111–123.
- [25] J.A. Brown, et al., Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA, *Proc. Natl. Acad. Sci. U. S. A.* 113 (49) (2016) 14013–14018.

- [26] D. Bernard, et al., A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression, *EMBO J.* 29 (18) (2010) 3082–3093.
- [27] R. Miyagawa, et al., Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles, *RNA* 18 (4) (2012) 738–751.
- [28] D.L. Spector, A.I. Lamond, Nuclear speckles, *Cold Spring Harb. Perspect. Biol.* 3 (2) (2011).
- [29] A.I. Lamond, D.L. Spector, Nuclear speckles: a model for nuclear organelles, *Nat. Rev. Mol. Cell Biol.* 4 (8) (2003) 605–612.
- [30] L. Galganski, M.O. Urbanek, W.J. Krzyzosiak, Nuclear speckles: molecular organization, biological function and role in disease, *Nucleic Acids Res.* 45 (18) (2017) 10350–10368.
- [31] C.M. Clemson, et al., An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles, *Mol. Cell* 33 (6) (2009) 717–726.
- [32] V. Tripathi, et al., The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation, *Mol. Cell* 39 (6) (2010) 925–938.
- [33] S. Nakagawa, et al., Malat1 is not an essential component of nuclear speckles in mice, *RNA* 18 (8) (2012) 1487–1499.
- [34] M. Eissmann, et al., Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development, *RNA Biol.* 9 (8) (2012) 1076–1087.
- [35] J.A. West, et al., The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites, *Mol. Cell* 55 (5) (2014) 791–802.
- [36] B. Li, et al., Activation of LTPB3 gene by a long noncoding RNA (lncRNA) MALAT1 transcript in mesenchymal stem cells from multiple myeloma, *J. Biol. Chem.* 289 (42) (2014) 29365–29375.
- [37] L. Yang, et al., ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs, *Cell* 147 (4) (2011) 773–788.
- [38] E. Leucci, et al., microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus, *Sci. Rep.* 3 (2013) 2535.
- [39] M. Zhuang, et al., MALAT1 sponges miR-106b-5p to promote the invasion and metastasis of colorectal cancer via SLAIN2 enhanced microtubules mobility, *EBioMedicine* 41 (2019) 286–298.
- [40] F. Li, et al., MALAT1 regulates miR-34a expression in melanoma cells, *Cell Death Dis.* 10 (6) (2019) 389.
- [41] F. Tao, et al., miR-211 sponges lncRNA MALAT1 to suppress tumor growth and progression through inhibiting PPH19 in ovarian carcinoma, *Faseb. J.* (2018), fj201800495RR.
- [42] M. Zhao, et al., MALAT1: a long non-coding RNA highly associated with human cancers, *Oncol. Lett.* 16 (1) (2018) 19–26.
- [43] Z.X. Li, et al., MALAT1: a potential biomarker in cancer, *Canc. Manag. Res.* 10 (2018) 6757–6768.
- [44] A. Bhan, M. Soleimani, S.S. Mandal, Long noncoding RNA and cancer: a new paradigm, *Canc. Res.* 77 (15) (2017) 3965–3981.
- [45] A. Sanchez Calle, et al., Emerging roles of long non-coding RNA in cancer, *Canc. Sci.* 109 (7) (2018) 2093–2100.
- [46] T. Gutschner, et al., The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells, *Canc. Res.* 73 (3) (2013) 1180–1189.
- [47] Q. Ji, et al., Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex, *Br. J. Canc.* 111 (4) (2014) 736–748.
- [48] P.J. McCown, et al., Secondary structural model of human MALAT1 reveals multiple structure-function relationships, *Int. J. Mol. Sci.* 20 (22) (2019).
- [49] M.C. Wang, et al., Secondary structural model of MALAT1 becomes unstructured in chronic myeloid leukemia and undergoes structural rearrangement in cervical cancer, *Noncoding RNA* 7 (1) (2021).
- [50] L. Lipovich, et al., Activity-dependent human brain coding/noncoding gene regulatory networks, *Genetics* 192 (3) (2012) 1133–1148.
- [51] K.M. Michalik, et al., Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth, *Circ. Res.* 114 (9) (2014) 1389–1397.
- [52] M. Vausort, D.R. Wagner, Y. Devaux, Long noncoding RNAs in patients with acute myocardial infarction, *Circ. Res.* 115 (7) (2014) 668–677.
- [53] Z.H. Zhao, et al., Long non-coding RNA MALAT1 functions as a mediator in cardioprotective effects of fentanyl in myocardial ischemia-reperfusion injury, *Cell Biol. Int.* 41 (1) (2017) 62–70.
- [54] J. Zhang, et al., Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia, *Exp. Neurol.* 277 (2016) 162–170.
- [55] J. Yao, et al., Long non-coding RNA MALAT1 regulates retinal neurodegeneration through CREB signaling, *EMBO Mol. Med.* 8 (4) (2016) 346–362.
- [56] Q.S. Zhang, et al., Beta-asarone protects against MPTP-induced Parkinson's disease via regulating long non-coding RNA MALAT1 and inhibiting alpha-synuclein protein expression, *Biomed. Pharmacother.* 83 (2016) 153–159.
- [57] H. Ren, et al., Association of circulating long non-coding RNA MALAT1 in diagnosis, disease surveillance, and prognosis of acute ischemic stroke, *Braz. J. Med. Biol. Res.* 53 (12) (2020), e9174.
- [58] F. Prinz, et al., The implications of the long non-coding RNA NEAT1 in non-cancerous diseases, *Int. J. Mol. Sci.* 20 (3) (2019).
- [59] Z. Wang, K. Li, W. Huang, Long non-coding RNA NEAT1-centric gene regulation, *Cell. Mol. Life Sci.* 77 (19) (2020) 3769–3779.
- [60] Y.T. Sasaki, et al., MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles, *Proc. Natl. Acad. Sci. U. S. A.* 106 (8) (2009) 2525–2530.
- [61] S. Nakagawa, T. Yamazaki, T. Hirose, Molecular dissection of nuclear paraspeckles: towards understanding the emerging world of the RNP milieu, *Open Biol.* 8 (10) (2018).
- [62] A.H. Fox, A.I. Lamond, Paraspeckles, *Cold Spring Harb. Perspect. Biol.* 2 (7) (2010) a000687.
- [63] G. Pisani, B. Baron, Nuclear paraspeckles function in mediating gene regulatory and apoptotic pathways, *Noncoding RNA Res.* 4 (4) (2019) 128–134.
- [64] A.H. Fox, et al., Paraspeckles: where long noncoding RNA meets phase separation, *Trends Biochem. Sci.* 43 (2) (2018) 124–135.
- [65] H. Sunwoo, et al., MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles, *Genome Res.* 19 (3) (2009) 347–359.
- [66] L.L. Chen, G.G. Carmichael, Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA, *Mol. Cell* 35 (4) (2009) 467–478.
- [67] K.V. Prasanth, et al., Regulating gene expression through RNA nuclear retention, *Cell* 123 (2) (2005) 249–263.
- [68] Y. Wang, et al., Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria, *Nat. Cell Biol.* 20 (10) (2018) 1145–1158.
- [69] Z. Huang, et al., Depletion of lncRNA NEAT1 rescues mitochondrial dysfunction through NEDD4L-dependent PINK1 degradation in animal models of Alzheimer's disease, *Front. Cell. Neurosci.* 14 (2020) 28.
- [70] L. Jiang, et al., NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing, *Nat. Struct. Mol. Biol.* 24 (10) (2017) 816–824.
- [71] D.R. Cooper, et al., Long non-coding RNA NEAT1 associates with SRP40 to temporally regulate PPARgamma2 splicing during adipogenesis in 3T3-L1 cells, *Genes* 5 (4) (2014) 1050–1063.
- [72] Z. Wang, et al., NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription, *Cell. Mol. Life Sci.* 74 (6) (2017) 1117–1131.
- [73] Z. Wang, et al., NEAT1 regulates neuroglial cell mediating Abeta clearance via the epigenetic regulation of endocytosis-related genes expression, *Cell. Mol. Life Sci.* 76 (15) (2019) 3005–3018.
- [74] W. Shen, et al., Depletion of NEAT1 lncRNA attenuates nucleolar stress by releasing sequestered P54nrb and PSF to facilitate c-Myc translation, *PLoS One* 12 (3) (2017), e0173494.
- [75] X. Shui, et al., Knockdown of lncRNA NEAT1 inhibits Th17/CD4(+) T cell differentiation through reducing the STAT3 protein level, *J. Cell. Physiol.* 234 (12) (2019) 22477–22484.
- [76] M. Zhang, et al., The lncRNA NEAT1 activates Wnt/beta-catenin signaling and promotes colorectal cancer progression via interacting with DDX5, *J. Hematol. Oncol.* 11 (1) (2018) 113.
- [77] X. Tan, et al., Knockdown of lncRNA NEAT1 suppresses hypoxia-induced migration, invasion and glycolysis in anaplastic thyroid carcinoma cells through regulation of miR-206 and miR-599, *Canc. Cell Int.* 20 (2020) 132.
- [78] S. Li, et al., Long non-coding RNA NEAT1 promotes the proliferation, migration, and metastasis of human breast-cancer cells by inhibiting miR-146b-5p expression, *Canc. Manag. Res.* 12 (2020) 6091–6101.
- [79] M. Luo, et al., Long noncoding RNA NEAT1 promotes ovarian cancer cell invasion and migration by interacting with miR1321 and regulating tight junction protein 3 expression, *Mol. Med. Rep.* 22 (4) (2020) 3429–3439.
- [80] M. Gao, et al., Long non-coding RNA NEAT1 serves as sponge for miR-365a-3p to promote gastric cancer progression via regulating ABCA4, *Oncotargets Ther.* 13 (2020) 3977–3985.
- [81] Z. Guo, et al., Long non-coding RNA-NEAT1, a sponge for miR-98-5p, promotes expression of oncogene HMG2 in prostate cancer, *Biosci. Rep.* 39 (9) (2019).
- [82] Y. Li, et al., lncRNA NEAT1 regulates cell viability and invasion in esophageal squamous cell carcinoma through the miR-129/CTBP2 Axis, *Dis. Markers* 2017 (2017) 5314649.
- [83] W. Yan, et al., lncRNA NEAT1 promotes autophagy in MPTP-induced Parkinson's disease through stabilizing PINK1 protein, *Biochem. Biophys. Res. Commun.* 496 (4) (2018) 1019–1024.
- [84] Y. Liu, Z. Lu, Long non-coding RNA NEAT1 mediates the toxic of Parkinson's disease induced by MPTP/MPP+ via regulation of gene expression, *Clin. Exp. Pharmacol. Physiol.* 45 (8) (2018) 841–848.
- [85] Y. Nishimoto, et al., The long non-coding RNA nuclear-enriched abundant transcript 1.2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis, *Mol. Brain* 6 (2013) 31.
- [86] J.S. Sunwoo, et al., Altered expression of the long noncoding RNA NEAT1 in Huntington's disease, *Mol. Neurobiol.* 54 (2) (2017) 1577–1586.
- [87] F.T. Bu, et al., lncRNA NEAT1: shedding light on mechanisms and opportunities in liver diseases, *Liver Int.* 40 (11) (2020) 2612–2626.
- [88] X. Wang, Down-regulation of lncRNA-NEAT1 alleviated the non-alcoholic fatty liver disease via mTOR/S6K1 signaling pathway, *J. Cell. Biochem.* 119 (2) (2018) 1567–1574.
- [89] Y.Q. Soh, et al., Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes, *Cell* 159 (4) (2014) 800–813.
- [90] J.L. Mueller, et al., Independent specialization of the human and mouse X chromosomes for the male germ line, *Nat. Genet.* 45 (9) (2013) 1083–1087.
- [91] I. Pinheiro, E. Heard, X chromosome inactivation: new players in the initiation of gene silencing, *F1000Research* 6 (2017).
- [92] A. Loda, E. Heard, Xist RNA in action: past, present, and future, *PLoS Genet.* 15 (9) (2019), e1008333.

- [93] R. Galupa, E. Heard, X-chromosome inactivation: a crossroads between chromosome architecture and gene regulation, *Annu. Rev. Genet.* 52 (2018) 535–566.
- [94] N. Brockdorff, J.S. Bowness, G. Wei, Progress toward understanding chromosome silencing by Xist RNA, *Genes Dev.* 34 (11–12) (2020) 733–744.
- [95] H. Fang, C.M. Disteche, J.B. Berletch, X inactivation and escape: epigenetic and structural features, *Front. Cell. Dev. Biol.* 7 (2019) 219.
- [96] N. Brockdorff, Localized accumulation of Xist RNA in X chromosome inactivation, *Open Biol.* 9 (12) (2019) 190213.
- [97] E.B. Hook, D. Warburton, The distribution of chromosomal genotypes associated with Turner's syndrome: livebirth prevalence rates and evidence for diminished fetal mortality and severity in genotypes associated with structural X abnormalities or mosaicism, *Hum. Genet.* 64 (1) (1983) 24–27.
- [98] A. Cockwell, et al., A cytogenetic and molecular study of a series of 45,X fetuses and their parents, *J. Med. Genet.* 28 (3) (1991) 151–155.
- [99] N. Brockdorff, et al., Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome, *Nature* 351 (6324) (1991) 329–331.
- [100] G. Borsani, et al., Characterization of a murine gene expressed from the inactive X chromosome, *Nature* 351 (6324) (1991) 325–329.
- [101] S.D. Brown, XIST and the mapping of the X chromosome inactivation centre, *Bioessays* 13 (11) (1991) 607–612.
- [102] N. Brockdorff, et al., The product of the mouse xist gene is a 15 Kb inactive X-specific transcript containing No conserved Orf and located in the nucleus, *Cell* 71 (3) (1992) 515–526.
- [103] C.J. Brown, et al., The human xist gene - analysis of a 17 Kb inactive X-specific rna that contains conserved repeats and is highly localized within the nucleus, *Cell* 71 (3) (1992) 527–542.
- [104] A. Wutz, T.P. Rasmussen, R. Jaenisch, Chromosomal silencing and localization are mediated by different domains of Xist RNA, *Nat. Genet.* 30 (2) (2002) 167–174.
- [105] Y. Hoki, et al., A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse, *Development* 136 (1) (2009) 139–146.
- [106] D.P. Patil, et al., m(6)A RNA methylation promotes XIST-mediated transcriptional repression, *Nature* 537 (7620) (2016) 369–373.
- [107] B. Moindrot, et al., A pooled shRNA screen identifies Rbm15, spen, and Wtap as factors required for xist RNA-mediated silencing, *Cell Rep.* 12 (4) (2015) 562–572.
- [108] C.A. McHugh, et al., The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3, *Nature* 521 (7551) (2015) 232–236.
- [109] A. Minajigi, et al., Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation, *Science* 349 (6245) (2015).
- [110] C. Chu, et al., Systematic discovery of Xist RNA binding proteins, *Cell* 161 (2) (2015) 404–416.
- [111] A. Monfort, et al., Identification of spen as a crucial factor for xist function through forward genetic screening in haploid embryonic stem cells, *Cell Rep.* 12 (4) (2015) 554–561.
- [112] M.G. Guenther, O. Barak, M.A. Lazar, The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3, *Mol. Cell Biol.* 21 (18) (2001) 6091–6101.
- [113] S.H. You, et al., Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo, *Nat. Struct. Mol. Biol.* 20 (2) (2013) 182–187.
- [114] Y. Shi, et al., Sharp, an inducible cofactor that integrates nuclear receptor repression and activation, *Genes Dev.* 15 (9) (2001) 1140–1151.
- [115] D. Cirillo, et al., Quantitative predictions of protein interactions with long noncoding RNAs, *Nat. Methods* 14 (1) (2016) 5–6.
- [116] C.K. Chen, et al., Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing, *Science* 354 (6311) (2016) 468–472.
- [117] G. Pintacuda, et al., hnRNPK recruits PCGF3/5-PRC1 to the xist RNA B-repeat to establish polycomb-mediated chromosomal silencing, *Mol. Cell* 68 (5) (2017) 955–969 e10.
- [118] D. Colognori, et al., Xist deletion analysis reveals an interdependency between xist RNA and polycomb complexes for spreading along the inactive X, *Mol. Cell* 74 (1) (2019) 101–117 e10.
- [119] N. Yamada, et al., Xist exon 7 contributes to the stable localization of xist RNA on the inactive X-chromosome, *PLoS Genet.* 11 (8) (2015), e1005430.
- [120] D. Smeets, et al., Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci, *Epigenet. Chromatin* 7 (2014) 8.
- [121] Y. Hasegawa, et al., The matrix protein hnRNP U is required for chromosomal localization of Xist RNA, *Dev. Cell* 19 (3) (2010) 469–476.
- [122] H.J. Kolpa, F.O. Fackelmayer, J.B. Lawrence, SAF-A requirement in anchoring XIST RNA to chromatin varies in transformed and primary cells, *Dev. Cell* 39 (1) (2016) 9–10.
- [123] T. Sakaguchi, et al., Control of chromosomal localization of xist by hnRNP U family molecules, *Dev. Cell* 39 (1) (2016) 11–12.
- [124] Y. Jeon, J.T. Lee, YY1 tethers Xist RNA to the inactive X nucleation center, *Cell* 146 (1) (2011) 119–133.
- [125] Q. Lv, et al., D-repeat in the XIST gene is required for X chromosome inactivation, *RNA Biol.* 13 (2) (2016) 172–176.
- [126] J. Chaumeil, et al., A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced, *Genes Dev.* 20 (16) (2006) 2223–2237.
- [127] N. Jansz, et al., SmcHD1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters, *Nat. Struct. Mol. Biol.* 25 (9) (2018) 766–777.
- [128] C.Y. Wang, et al., SMCHD1 Merges chromosome compartments and assists formation of super-structures on the inactive X, *Cell* 174 (2) (2018) 406–421 e25.
- [129] M.R. Gdula, et al., The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome, *Nat. Commun.* 10 (1) (2019) 30.
- [130] M.E. Blewitt, et al., SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation, *Nat. Genet.* 40 (5) (2008) 663–669.
- [131] A.V. Gendrel, et al., SmcHD1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome, *Dev. Cell* 23 (2) (2012) 265–279.
- [132] M. Denholtz, et al., Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization, *Cell Stem Cell* 13 (5) (2013) 602–616.
- [133] S. Kundu, et al., Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation, *Mol. Cell* 71 (1) (2018) 191.
- [134] S. Schoenfelder, et al., Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome, *Nat. Genet.* 47 (10) (2015) 1179–1186.
- [135] Y. Du, et al., LncRNA XIST acts as a tumor suppressor in prostate cancer through sponging miR-23a to modulate RKIP expression, *Oncotarget* 8 (55) (2017) 94358–94370.
- [136] W. Wei, et al., LncRNA XIST promotes pancreatic cancer proliferation through miR-133a/EGFR, *J. Cell. Biochem.* 118 (10) (2017) 3349–3358.
- [137] Y.L. Zhang, et al., The lncRNA XIST exhibits oncogenic properties via regulation of miR-449a and Bcl-2 in human non-small cell lung cancer, *Acta Pharmacol. Sin.* 38 (3) (2017) 371–381.
- [138] D. Yue, et al., Silencing of long noncoding RNA XIST attenuated Alzheimer's disease-related BACE1 alteration through miR-124, *Cell Biol. Int.* 44 (2) (2020) 630–636.
- [139] U. Braunschweig, et al., Dynamic integration of splicing within gene regulatory pathways, *Cell* 152 (6) (2013) 1252–1269.
- [140] T.W. Nilsen, B.R. Graveley, Expansion of the eukaryotic proteome by alternative splicing, *Nature* 463 (7280) (2010) 457–463.
- [141] M.C. Wahl, C.L. Will, R. Luhrmann, The spliceosome: design principles of a dynamic RNP machine, *Cell* 136 (4) (2009) 701–718.
- [142] L.L. Chen, L. Yang, Regulation of circRNA biogenesis, *RNA Biol.* 12 (4) (2015) 381–388.
- [143] C. Huang, G. Shan, What happens at or after transcription: insights into circRNA biogenesis and function, *Transcription* 6 (4) (2015) 61–64.
- [144] Z. Song, et al., Antisense oligonucleotide technology can be used to investigate a circular but not linear RNA-mediated function for its encoded gene locus, *Sci. China Life Sci.* (2020), <https://doi.org/10.1007/s11427-020-1743-8>.
- [145] L. Gao, et al., Circular RNAs from BOULE play conserved roles in protection against stress-induced fertility decline, *Sci. Adv.* 6 (46) (2020).
- [146] H.L. Sanger, et al., Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures, *Proc. Natl. Acad. Sci. U. S. A.* 73 (11) (1976) 3852–3856.
- [147] T.B. Hansen, et al., Natural RNA circles function as efficient microRNA sponges, *Nature* 495 (7441) (2013) 384–388.
- [148] S. Memczak, et al., Circular RNAs are a large class of animal RNAs with regulatory potency, *Nature* 495 (7441) (2013) 333–338.
- [149] L.S. Kristensen, et al., The biogenesis, biology and characterization of circular RNAs, *Nat. Rev. Genet.* 20 (11) (2019) 675–691.
- [150] M. Zhou, et al., New progresses of circular RNA biology: from nuclear export to degradation, *RNA Biol.* (2020) 1–9.
- [151] L.L. Patop, S. Wust, S. Kadener, Past, present, and future of circRNAs, *EMBO J.* 38 (16) (2019), e100836.
- [152] L. Chen, G. Shan, CircRNA in cancer: Fundamental mechanism and clinical potential, *Canc. Lett.* 505 (2021) 49–57.
- [153] J. Li, et al., Circular RNAs in cancer: biogenesis, function, and clinical significance, *Trends Canc.* 6 (4) (2020) 319–336.
- [154] L.L. Chen, The expanding regulatory mechanisms and cellular functions of circular RNAs, *Nat. Rev. Mol. Cell Biol.* 21 (8) (2020) 475–490.
- [155] S.J. Conn, et al., The RNA binding protein quaking regulates formation of circRNAs, *Cell* 160 (6) (2015) 1125–1134.
- [156] R. Ashwal-Fluss, et al., circRNA biogenesis competes with pre-mRNA splicing, *Mol. Cell* 56 (1) (2014) 55–66.
- [157] A. Ivanov, et al., Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals, *Cell Rep.* 10 (2) (2015) 170–177.
- [158] M.A. Khan, et al., RBM20 regulates circular RNA production from the titin gene, *Circ. Res.* 119 (9) (2016) 996–1003.
- [159] L. Errichelli, et al., FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons, *Nat. Commun.* 8 (2017) 14741.
- [160] C.X. Liu, et al., Structure and degradation of circular RNAs regulate PKR activation in innate immunity, *Cell* 177 (4) (2019) 865–880 e21.
- [161] J.W. Fischer, et al., Structure-mediated RNA decay by UPF1 and G3BP1, *Mol. Cell* 78 (1) (2020) 70–84 e6.
- [162] R. Jia, et al., Defining an evolutionarily conserved role of GW182 in circular RNA degradation, *Cell Discov.* 5 (2019) 45.
- [163] O.H. Park, et al., Endoribonucleolytic cleavage of m(6)a-containing RNAs by RNase P/MRP complex, *Mol. Cell* 74 (3) (2019) 494–507 e8.
- [164] T.B. Hansen, et al., miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA, *EMBO J.* 30 (21) (2011) 4414–4422.

- [165] C. Huang, et al., A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs, *Genes Dev.* 32 (9–10) (2018) 639–644.
- [166] Z. Li, M.G. Kearse, C. Huang, The nuclear export of circular RNAs is primarily defined by their length, *RNA Biol.* 16 (1) (2019) 1–4.
- [167] Z. Li, et al., Exon-intron circular RNAs regulate transcription in the nucleus, *Nat. Struct. Mol. Biol.* 22 (3) (2015) 256–264.
- [168] Y. Zhang, et al., Circular intronic long noncoding RNAs, *Mol. Cell* 51 (6) (2013) 792–806.
- [169] L. Stoll, et al., A circular RNA generated from an intron of the insulin gene controls insulin secretion, *Nat. Commun.* 11 (1) (2020) 5611.
- [170] N. Chen, et al., A novel FLI1 exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1, *Genome Biol.* 19 (1) (2018) 218.
- [171] A. Bronisz, et al., The nuclear DICER-circular RNA complex drives the deregulation of the glioblastoma cell microRNAome, *Sci. Adv.* 6 (51) (2020).
- [172] P. Xia, et al., A circular RNA protects dormant hematopoietic stem cells from DNA sensor cGAS-mediated exhaustion, *Immunity* 48 (4) (2018) 688–701 e7.
- [173] V.M. Conn, et al., A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation, *Nat. Plants* 3 (2017) 17053.
- [174] Y. Liu, et al., Back-spliced RNA from retrotransposon binds to centromere and regulates centromeric chromatin loops in maize, *PLoS Biol.* 18 (1) (2020), e3000582.
- [175] O.Z. Aydin, W. Vermeulen, H. Lans, ISWI chromatin remodeling complexes in the DNA damage response, *Cell Cycle* 13 (19) (2014) 3016–3025.
- [176] F. Erdel, K. Rippe, Binding kinetics of human ISWI chromatin-remodelers to DNA repair sites elucidate their target location mechanism, *Nucleus* 2 (2) (2011) 105–112.
- [177] X. Xu, et al., CircRNA inhibits DNA damage repair by interacting with host gene, *Mol. Canc.* 19 (1) (2020) 128.
- [178] Y. Long, et al., How do lncRNAs regulate transcription? *Sci. Adv.* 3 (9) (2017), eaao2110.
- [179] Q. Sun, Q. Hao, K.V. Prasanth, Nuclear long noncoding RNAs: key regulators of gene expression, *Trends Genet.* 34 (2) (2018) 142–157.
- [180] C.Y. Guh, Y.H. Hsieh, H.P. Chu, Functions and properties of nuclear lncRNAs from systematically mapping the interactomes of lncRNAs, *J. Biomed. Sci.* 27 (1) (2020) 44.
- [181] C. Chu, et al., Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions, *Mol. Cell* 44 (4) (2011) 667–678.
- [182] H.P. Chu, et al., TERRA RNA antagonizes ATRX and protects telomeres, *Cell* 170 (1) (2017) 86–101, e16.
- [183] M.R. Mumbach, et al., HiChIRP reveals RNA-associated chromosome conformation, *Nat. Methods* 16 (6) (2019) 489–492.
- [184] M.D. Simon, et al., High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation, *Nature* 504 (7480) (2013) 465–469.
- [185] M.D. Simon, et al., The genomic binding sites of a noncoding RNA, *Proc. Natl. Acad. Sci. U. S. A.* 108 (51) (2011) 20497–20502.
- [186] L. Chen, et al., R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters, *Mol. Cell* 68 (4) (2017) 745–757 e5.
- [187] J.G. Dumelie, S.R. Jaffrey, Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq, *eLife* 6 (2017).
- [188] P.A. Ginno, et al., R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters, *Mol. Cell* 45 (6) (2012) 814–825.
- [189] L.A. Sanz, et al., Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals, *Mol. Cell* 63 (1) (2016) 167–178.
- [190] M. Hafner, et al., Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP, *Cell* 141 (1) (2010) 129–141.
- [191] J. Konig, et al., iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution, *Nat. Struct. Mol. Biol.* 17 (7) (2010) 909–915.
- [192] J. Ule, et al., CLIP identifies Nova-regulated RNA networks in the brain, *Science* 302 (5648) (2003) 1212–1215.
- [193] Y. Sugimoto, et al., hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1, *Nature* 519 (7544) (2015) 491–494.
- [194] Y. Ding, et al., In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features, *Nature* 505 (7485) (2014) 696–700.
- [195] D. Loughrey, et al., SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing, *Nucleic Acids Res.* 42 (21) (2014).
- [196] J.B. Lucks, et al., Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq), *Proc. Natl. Acad. Sci. U. S. A.* 108 (27) (2011) 11063–11068.
- [197] R.C. Spitale, et al., Structural imprints in vivo decode RNA regulatory mechanisms, *Nature* 519 (7544) (2015) 486–490.
- [198] J.G. Underwood, et al., FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing, *Nat. Methods* 7 (12) (2010) 995–1001.
- [199] A.R. Bassett, et al., Considerations when investigating lncRNA function in vivo, *eLife* 3 (2014), e03058.