

Chapter 2

Long Noncoding RNA: Genome Organization and Mechanism of Action

Vijay Suresh Akhade, Debosree Pal, and Chandrasekhar Kanduri

Abstract For the last four decades, we have known that noncoding RNAs maintain critical housekeeping functions such as transcription, RNA processing, and translation. However, in the late 1990s and early 2000s, the advent of high-throughput sequencing technologies and computational tools to analyze these large sequencing datasets facilitated the discovery of thousands of small and long noncoding RNAs (lncRNAs) and their functional role in diverse biological functions. For example, lncRNAs have been shown to regulate dosage compensation, genomic imprinting, pluripotency, cell differentiation and development, immune response, etc. Here we review how lncRNAs bring about such copious functions by employing diverse mechanisms such as translational inhibition, mRNA degradation, RNA decoys, facilitating recruitment of chromatin modifiers, regulation of protein activity, regulating the availability of miRNAs by sponging mechanism, etc. In addition, we provide a detailed account of different mechanisms as well as general principles by which lncRNAs organize functionally different nuclear sub-compartments and their impact on nuclear architecture.

Keywords Long noncoding RNA • Chromatin • Gene regulation • Genome organization • Epigenetics

V.S. Akhade • C. Kanduri (✉)

Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg 40530, Sweden
e-mail: kanduri.chandrasekhar@gu.se

D. Pal

Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O, Bangalore 560064, India

2.1 Introduction

Noncoding RNA (ncRNA), once thought as a part of transcriptional noise, now represents a novel regulatory layer in the transcriptional and posttranscriptional gene regulation. Their rise from transcriptional noise to novel regulators of gene expression is well supported by their documented functional roles in various aspects of gene regulation, including epigenetic regulation, X chromosome inactivation, genomic imprinting, nuclear and cytoplasmic trafficking, transcription, and mRNA splicing [1, 2]. Latest genome-wide annotation studies based on high-throughput transcriptomics from single cell embryo to differentiated tissue cell types reveal that more than two-thirds of the mammalian genome is transcribed encoding tens of thousands of different classes of small and long noncoding RNA (lncRNAs). lncRNAs represent the largest class among the noncoding RNA subtypes, and according to the latest estimates, there are about more than 58,084 transcripts, which outnumber the protein-coding RNAs. lncRNAs have also emerged as key regulators in a wide range of biological processes such as cell proliferation, cell cycle, metabolism, apoptosis, differentiation and maintenance of pluripotency, etc. [3, 4]. They have not only shown to typically regulate gene regulation, but lncRNAs employ diverse mechanisms for their function, and these would be described below in detail.

2.2 lncRNA and Chromatin Regulation

lncRNAs have emerged as crucial players in the regulation of transcription via modulation of chromatin [5–9]. lncRNAs regulate chromatin structure at different functional steps which include histone modifications, DNA methylation, and chromatin remodeling. Execution of this function by lncRNAs across a broad evolutionary spectrum is suggestive of their conserved role in chromatin regulation despite lack of primary sequence conservation.

2.3 lncRNAs Regulate Chromatin Structure via Histone Modifications

Interaction of lncRNAs with histone-modifying complexes is very prominent with respect to two polycomb repressive complexes, PRC1 and, in particular, PRC2, which mediates methylation of lysine 27 on histone 3 (H3K27me), a histone mark associated with poised or repressed transcriptional status. A general mechanistic representation of regulation of histone modifications by lncRNAs is depicted in Fig. 2.1a. The first documentation of lncRNA and PRC2 interaction was from the studies on mammalian X chromosome inactivation, where X-inactive specific

Mechanisms of lncRNA action

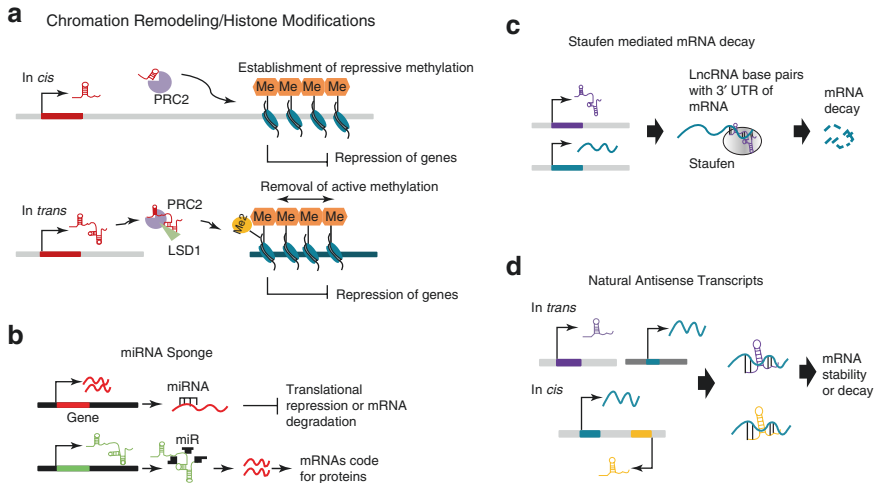


Fig. 2.1 Schematic representation of few mechanisms employed by lncRNAs in regulation of gene expression which include regulation of active and repressive histone marks (a), mRNA translation by acting as a miRNA sponge (b), Staufen-mediated decay of mRNA (c), and mRNA stability or decay by functioning as natural antisense transcripts (d)

transcript (*Xist*), a lncRNA that is highly expressed from the inactive X chromosomes in females (Xi), mediates the recruitment of PRC2 to the Xi to silence gene expression [10–12]. Another prominent example is HOTAIR lncRNA, which is transcribed from the *HoxC* gene cluster but mediates transcriptional gene silencing in the *HoxD* locus in trans via targeting PRC2 complex and the histone H3K4me1/2 demethylase LSD1 [13]. Recently, a murine lincRNA Pint (p53 induced noncoding transcript) was demonstrated to connect p53 activation with epigenetic silencing by PRC2 [14]. Pint controls cell survival and proliferation through regulating the TGF- β , MAPK, and p53 pathways. Pint mediates H3K27 trimethylation of target gene promoters via recruitment of PRC2. Also, the expression of PRC2 is required for the functional activity of Pint. Braveheart (Bvht), a heart associated lncRNA in mouse, is activated during early cardiac differentiation [15]. Bvht mediates epigenetic regulation of cardiac commitment through its interaction with SUZ12, a component of PRC2 complex during cardiomyocyte differentiation. Another similar example is the Fendrr lncRNA (*Foxf1* adjacent noncoding developmental regulatory RNA) which is required for heart and body wall development in mouse. Fendrr acts by modifying the chromatin signatures of its target genes through binding to both the PRC2 and TrxG/MLL complexes. Fendrr interacts with WDR5 which is a component of the MLL complexes that mediates H3K4 methylation associated with loci that are actively transcribed or primed for activation [16, 17]. lncRNA SRA (steroid receptor RNA activator) also participates in transcriptional regulation through complex formation with trithorax group (TrxG) and polycomb repressive complex 2 (PRC2) complexes [18]. Very recently, *HoxB* lincRNA is shown to be required for

HoxB gene activation and mesoderm specification (*HoxB* lincRNA cell reports reference). Mechanistically, *HoxB* lincRNA controls chromatin dynamics through the recruitment of Set1/MLL complexes. Interaction of WDR5 has been reported with more than 200 lincRNAs in mouse embryonic stem cells (mESCs) [19]. The importance of these interactions is reflected by the requirement of WDR5-lincRNA association for binding of WDR5 to chromatin. Taking into consideration more specific studies, two lincRNAs, HOTTIP and NeST, have been described to recruit WDR5 to their neighboring genes and thus enhance their transcription [20, 21].

Example of an lincRNA interacting with the PRC1 complex is FAL1 RNA (focally amplified lincRNA on chromosome 1) [22]. FAL1 interacts with BMI1, an essential subunit of PRC1. FAL1 not only regulates the protein stability of BMI1 but also regulates the association between BMI1 and the target promoter regions thereby modulating target gene expression. ANRIL, an lincRNA transcribed from the INK4B-ARF-INK4A tumor suppressor locus, was discovered in a family with inherited melanoma-neural system tumors [23]. ANRIL recruits CBX7 (a PRC1 component) via a POL II-dependent mechanism to its locus in order to repress the neighboring INK4B-ARF-INK4A genes [23, 24]. CBX7 binding to ANRIL contributes to CBX7 function, and disruption of CBX7-ANRIL interaction impacts the ability of CBX7 to repress the INK4b/ARF/INK4a locus and control senescence [24].

2.4 lincRNAs Regulate Chromatin Remodeling

In addition to interaction with histone-modifying enzymes and covalently modifying chromatin, lincRNAs also associate with chromatin remodeling complexes to regulate gene expression. The association of lincRNA *SChLAPI* (second chromosome locus associated with prostate-1) with the chromatin remodeling complex SWI/SNF in human prostate cancer cells is one such example [25]. Knockdown of lincRNA *SChLAPI* leads to impaired cell invasion and proliferation, while knockdown of SMARCB1 (component of SWI/SNF) promoted cancer progression, thus revealing their opposite functions in human prostate cancer cell lines. Also, *SChLAPI* interacts with the SNF5 subunit of the SWI/SNF complex and antagonizes the genome-wide localization and regulatory functions of the SWI/SNF complex, thereby resulting in genome-wide derepression of gene activity. Similarly, the lincRNA *Mhrt* (myosin heavy chain-associated RNA transcripts or Myheart) binds to BRG1, the ATPase subunit of the SWI/SNF complex, and prevents Brg1 from recognizing its genomic DNA targets [26]. Recently, a very interesting study comprehensively identified the *Xist* lincRNA interactome which comprised of cohesins, condensins, topoisomerases, RNA helicases, chromatin remodelers, and modifiers [27]. Stable knockdown of the components of SWI/SNF complex demonstrated that the *Xist*-SWI/SNF interaction is required for proper maintenance of PRC2 function on the inactive X chromosome (Xi). The *Evf2* lincRNA also interacts with SWI/SNF-related chromatin remodelers Brg1 and Baf170 [28, 29]. In contrast to

SChlAPI1, *Mhrt*, and *Xist*, *Eyf2* lncRNA increases the association of Brg1 with important regulatory enhancer regions but directly inhibits the ATPase and chromatin remodeling activities of Brg1.

lncRNAs and the SWI/SNF complex are not only shown in the context of gene repression but also reported to mediate gene activation. lncTCF7, which promotes the self-renewal of human liver cancer stem cells by activating Wnt signaling, interacts with and recruits the SWI/SNF complex to the TCF7 promoter thereby activating TCF7 expression and subsequently Wnt signaling [30]. linc-Cox2 is an important RNA for regulation of immune response genes, and its expression is highly induced in macrophages upon stimulation by LPS through the Toll-like receptors [31]. Recent studies have shed more light on the mechanism of transcriptional regulation by linc-Cox2. Upon LPS stimulation in macrophages, linc-Cox2 is assembled into the SWI/SNF complex and promotes the assembly of NF- κ B subunits into the SWI/SNF complex [32]. This ultimately leads to increased recruitment of SWI/SNF complex and transactivation of the late primary inflammatory response genes. linc-Cox2, also stimulated in response to TNF- α stimulation in murine intestinal epithelial cells, represses the transcription from the I12b promoter, a secondary late-responsive gene induced by TNF- α , through targeting Mi-2/NuRD repressor complex [33].

2.5 eRNAs and Higher-Order Chromatin Organization

Enhancer-derived lncRNAs (eRNAs) are also documented in the control of regulatory contacts between enhancers and the cognate promoter through chromosome looping [34, 35]. Until recently, eRNAs were considered primarily as by-products of transcription from enhancer-promoter interactions. However, with several eRNAs functionally implicated in diverse biological contexts, they are being considered as independent transcriptional units [34, 36, 37]. The classical enhancer RNAs (eRNAs) are short, unspliced transcripts that are a product of bidirectional transcription [34, 38]. The locus control region (LCR) of the β -globin cluster is the very first identified example of enhancer-linked transcripts [39, 40]. The eRNAs control gene expression, possibly by affecting looping between enhancers and promoters [41, 42]. For example, LUNAR1 is a T cell acute lymphoblastic leukemia (T-ALL)-specific lncRNA that is transcribed from the insulin-like growth factor 1 receptor (*IGF1R*) locus and that exhibits pro-oncogenic characteristics, such as stimulating T-ALL cell growth [43]. LUNAR1 itself activates the *IGF1R* locus in *cis* via chromosome looping, thus leading to sustained IGF1 signaling in T-ALL cells. Likewise, the lncRNAs PRNCR1 and PCGEM1 regulate androgen receptor-dependent gene activation programs by promoting enhancer-promoter looping in prostate cancer cells [44]. Also, there are lncRNAs with enhancer-like functions and are referred to as activating ncRNAs (ncRNA-a). These lncRNAs mediate DNA looping and chromatin enhancement via Mediator, a large transcriptional co-activating complex [45]. Two such eRNAs, ncRNA-a3 and ncRNA-a7, bring about the

recruitment and activation of Mediator complex. Activity of these eRNAs is compromised by the knockdown of Mediator components and conversely loss of the eRNAs led to the decreased recruitment of Mediator and Pol II to the target genes. The contacting of eRNAs by the Mediator subunits is essential for the chromosomal looping between the enhancer and the target gene, as demonstrated by chromosome conformation capture method. Recently, using mouse cortical neurons, it was shown that some eRNAs facilitate the release of the negative elongation factor (NELF), thus competitively derepressing paused Pol II and enabling productive elongation of the target RNA [46]. Similarly, eRNAs may also “trap” certain RNA-binding transcription factors at enhancers, thereby sustaining transcription factor-mediated regulation [47].

lncRNAs transcribed from enhancer regions can also have inhibitory effects on their target genes. The promoter deletion of the lncRNA *Haunt* (also known as *linc-Hoxa1*) leads to upregulation of several genes of the neighboring *HoxA* gene cluster [48]. A similar mechanism was also found for *Playrr* (D030025E07Rik), an lncRNA-encoded upstream of the homeodomain transcription factor *Pitx2* [49]. A CRISPR/Cas9-generated mutation resulting in *Playrr* RNA decay caused upregulation of *Pitx2* expression. An enhancer involved in *Pitx2* regulation in the gut overlaps the TSS of *Playrr*. The expression of *Playrr* interferes with the looping of the *Pitx2* promoter to its enhancer at the *Playrr* locus and thus affecting the *Pitx2* gene activation.

2.6 Long Noncoding RNAs in Genomic Imprinting

Genomic imprinting is a phenomenon in which only one of the alleles of an inherited parental pair is active, while the other one is maintained in an inactive state. This differential expression of the inherited parental alleles depends on the parent of origin; in some cases, an allele of a gene might be paternally imprinted, whereas in other cases it would be maternally imprinted. Imprinting is generally achieved by histone and/or DNA modification of the particular locus, and lately lncRNAs have been known to play a role in this phenomenon [50].

Airn lncRNA is 108 kb long, nuclear localized transcript, transcribed in an anti-sense direction from 3.7 kb imprinting control element (ICE) in intron 2 of the insulin-like growth factor type-2 receptor (*Igf2r*) gene [51]. It has been shown to control the parent of origin-specific expression of three genes: *Slc22a2*, *Slc22a3*, and *Igf2r*. Paternal but not the maternal inheritance of 3.7 kb ICE deletion leads to biallelic expression of all three genes, including *Igf2r*, and reduction in birth weight of mice [52]. A similar phenotype was also observed when the *Airn* RNA was prematurely truncated by the insertion of a polyadenylation signal 3 kb downstream of its transcription start site. By performing *Airn* RNA-TRAP experiments in mouse liver and placental cell types, it was observed that *Airn* interacts physically with the *Slc22a3*, which is situated >230 kb away from the *Airn* locus, and recruits G9a methyltransferase to the *Slc22a3* promoter on the paternal chromosome in cis

[53, 54]. Based on these observations, the authors have hypothesized that *Airn* might act in a manner similar to *XIST* wherein direct interaction of the RNA with its target DNA loci in *cis* on the paternal chromosome would lead to the silencing of *Scl22a3* through the enrichment of a repressive histone mark H3K9me3.

Kcnq1ot1 is another lncRNA whose role in genomic imprinting is well established. It is a 91 kb RNA encoded from 1 Mb *Kcnq1/Cdkn1c* locus on chromosome 7 that harbors several protein-coding genes [55, 56]. *Kcnq1ot1* is expressed from the paternal chromosome, whereas it is repressed on the maternal chromosome by CpG methylation. Its expression on the paternal chromosome correlates with repression of eight to ten neighboring protein-coding genes which span over megabase region (reviewed in [57]). Like *Xist* and *Airn*, *Kcnq1ot1* acts in *cis* on the paternal chromosome to silence *Kcnq1*, *Cdkn1c*, *Phlda2*, and *Slc22a18* genes in all tissues, i.e., embryonic and extraembryonic (ubiquitously imprinted loci) and *Osbp15*, *Tssc4*, and *Ascl2* genes only in the extraembryonic tissues (placental-specific imprinted loci). *Kcnq1ot1* also acts at the epigenetic level to impart silencing of the imprinted genes. *Kcnq1ot1* mediates silencing of neighboring genes on the paternal chromosome through establishing paternal allele-specific repressive histone modifications (H3K9me3 and H3K27me3) and DNA methylation at the target imprinted gene promoters by interacting with chromatin (PRC2 and G9a) and DNA (DNMT1) modification enzymes and guide them specifically to the target imprinted gene promoters in *cis* by acting as a scaffold. Previously, using episomal-based system coupled with cell culture experiments, 890 nucleotide functional RNA sequence was identified at the 5' end of the *Kcnq1ot1* RNA. Using transgenic mouse model (Δ 890 mice) [57], it has been shown that the 890 RNA sequence at the 5' end of *Kcnq1ot1* RNA is required for establishing the repressive histone modifications and DNA methylation at the promoters of ubiquitously imprinted genes *Kcnq1*, *Cdkn1c*, *Slc22a18*, and *Phlda2*. Interestingly, the Δ 890 mice showed a phenotype similar to *Dnmt*^{-/-} mice. On further analysis, it was seen that DNA methylation at the paternal *Cdkn1c* and *Slc22a18* loci was significantly reduced in the Δ 890 mice. Surprisingly, *Kcnq1* and *Phlda2* did not show any changes in DNA methylation patterns in the wild-type or the Δ 890 mice implying that their methylation status could be established at a very early stage of development and then transmitted as a memory. Thus *Kcnq1ot1* lncRNA acts through two independent mechanisms: either by recruiting heterochromatic chromatin remodelers or by DNA methyltransferases specifically to imprinted genes in a locus- and tissue-specific manner [58].

In mouse chromosome 7, there exists a whole cluster of genes covering around 600 kb on the genome which undergo the phenomenon of imprinting. *H19* and *Igf2* are two genes belonging to this cluster; corresponding to the human locus 11p15.5, *H19* is expressed maternally, whereas *Igf2* is expressed paternally [59, 60]. This imprinted locus, also known as imprinted gene network (IGN), is regulated in expression by the *H19* gene, and the network is itself involved in the development of the embryo. *H19* encodes for a 2.3 kb lncRNA, and it has been shown that targeted deletion of the gene induces an overgrowth phenotype and relieves imprinting on *Igf2* and several other genes in the IGN [59, 61]. Monnier et al. addressed the mechanism of action of *H19* and performed RNA immunoprecipitation experiments

wherein they found that H19 interacts with methyl-CpG-binding domain protein 1 (MBD1) [60]. MBD1 binds to methylated DNA to recruit various histone deacetylases and lysine methyl transferases that impart histone modifications leading to gene silencing. Indeed it was observed that in MBD1 silenced cells, five of the IGN genes, i.e., *Igf2*, *Slc38a4*, *Dcn*, *Dlk1*, and *Peg1* underwent derepression, whereas *Gtl2*, *Cdkn1c*, and *Igf2r* did not thereby implying that only for these five genes interplay of H19 and MBD1 is important for the imprinting to occur. Further ChIP experiments revealed that MBD1 binds directly to the DNA methylated regions (DMRs) of *Igf2*, *Slc38a4*, and *Peg1* and this interaction is hindered in the absence of H19. It was also observed that when H19 is expressed in trans in cells that harbor a deletion of the H19 transcription unit, *Igf2* undergoes a loss of its imprinting status. Together these studies stated that H19 interacts with MBD1 to maintain transcriptional repression and imprinting of the above genes through imparting repressive histone modifications such as H3K9me3 at these loci.

2.7 Posttranscriptional Regulation by lncRNAs

lncRNAs as sponges: An interesting mechanism of action of lncRNAs that has come into focus is their activity as microRNA (miRNA) sponges (Fig. 2.1b). lncRNAs that are known to exert such an action harbor sequences complementary to miRNA sequences thereby sequestering them and preventing them from binding to their targets. Such lncRNAs can arise from pseudogenes or be in the form of circular RNAs or be common intergenic lncRNAs possessing miRNA binding sites. Cesana et al. in 2011 have studied for the first time one such lncRNA *linc-MD1*, a muscle-specific lncRNA implicated in muscle differentiation [62]. Bioinformatic analysis showed the presence of two binding sites for miR-135 and one binding site for miR-133 on *linc-MD1*. Interestingly, both miRNAs target important transcription factors regulating myogenic differentiation: miR-135 targets MEF2C and miR-133 targets MAML1. Luciferase assay experiments, with the wild-type *linc-MD1* sequence or its mutant derivatives containing mutations on the respective miR binding sites, cloned downstream of the luciferase gene and revealed that upon miR overexpression, luciferase activity for the wild-type MD1 sequence is depleted but not for the mutant derivatives. This suggests that it is sequestration of miR-133 and miR-135 by *linc-MD1* that acts to maintain the expression of MAML1 and MEF2C during myogenic differentiation.

PTENP1 is a pseudogene that shares a high similarity with *PTEN*, its 3'-UTR being 1 kb shorter than *PTEN* itself [63]. *PTENP1* possesses a binding site for miR-499-5p which can target both *PTEN* and *PTENP1*. In a high-fat-diet-fed mice, it was observed that *PTENP1* levels were upregulated with a reduction in Akt/GSK signaling pathway and decreased glycogen contents. When high-fat-diet-fed mice were injected with shRNA against *PTENP1*, miR-499-5p levels did not change although Akt/GSK phosphorylation increased as did glycogen synthesis indicating a shift from the insulin resistance. However, the levels of *PTEN*, a target of miR-499-5p, were found to be significantly reduced that led to the activation of the Akt/GSK

signaling pathway. These studies proposed a mechanism whereby *PTENP1* acts as a decoy for miR-499-5p to regulate the activity of *PTEN* in insulin resistance.

Studies by Liang et al. directed toward understanding the role of lncRNAs in osteogenesis revealed the involvement of *H19* in promoting osteoblast differentiation from human mesenchymal stem cells (hMSCs) in vitro [64]. Bioinformatic analyses showed that the *H19* sequence harbor binding sites for miR-141 and miR-22; however, their ectopic expression in cells did not affect the *H19* levels. Subsequent luciferase reporter assays however showed a reduction in luciferase expression, and ablation of this effect when the miR binding sites were mutated thereby establishing that these miRs act through translational repression rather than degradation of the target. With respect to osteoblast differentiation, it was observed that miR-141 and miR-22 were both downregulated during in vitro differentiation, and ectopic expression of miRNA mimics of miR-141 and miR-22 prevented proper osteogenic differentiation of hMSCs. Furthermore, β -catenin was validated as a common target of both these miRNAs. Interestingly, Wnt signaling plays a significant role during osteogenesis. When a luciferase vector containing the miR-141 and miR-22 binding sites were coexpressed with a *H19* overexpression vector, the luciferase activity was found to be upregulated, suggesting that *H19* is probably involved in decoying the miRNAs. *H19* was also shown to activate Wnt signaling and increase β -catenin levels in cells. These experiments indicated that *H19* may act as a sponge for miR-141 and miR-22 which otherwise act as negative regulators of Wnt signaling, thereby causing Wnt-mediated osteogenic differentiation of hMSCs.

In a recent investigation, an interplay between *BC032469*, *hTERT*, and miR-1207-5p and miR-1266 was investigated in the proliferation of gastric cancer (GC) cells [65]. It was found to be among the most differentially expressed panel of lncRNAs perturbed in hTERT positive versus negative gastric cancer tissues. hTERT is the rate-limiting subunit of telomerase, and its elevated expression is associated with several malignancies. Knockdown of *BC032469* in gastric cancer cell lines results in significant downregulation of *hTERT*, and it also had a negative effect on the cell proliferation. It was shown that miR-1207-5p and miR-1266 can target both *hTERT* and *BC032469*. *BC032469* knockdown resulted in *hTERT* downregulation, and this could be partially reversed by providing anti miR-1207-5p. Overexpression of miR-1207-5p attenuated the *BC032469*-mediated hTERT activation. These results indicate that lncRNA *BC032469* acts a miRNA sponge to sequester miR-1207-5p, which otherwise binds to and degrades its target hTERT, leading to upregulation of hTERT activity in GC that contributes to growth and progression of gastric cancer.

2.8 lncRNAs in Staufen-Mediated mRNA Decay

The half-life of an mRNA is vital to cell survival and disease manifestation in eukaryotes. In the cytoplasm, mRNA stability by nonsense-mediated mRNA decay (NMD), nonstop mRNA decay (NSD), and no-go mRNA decay (NGD) [66]. Cis-regulatory elements in the 3'UTR of mRNAs also regulate their stability, one such

mechanism being through Staufen-mediated decay (SMD). STAU1 is an established protein effector of SMD, which binds to double-stranded RNA formed either by inter- or intramolecular base pairing. In the scenario of intermolecular base pairing, lncRNAs are known to play a function. Such lncRNA contains Alu elements within their sequences that base pair via partial complementarity to the 3'UTRs of target mRNAs, thereby activating STAU-mediated decay of mRNAs (Fig. 2.1c).

Studies by Gong and Maquat were focused on the identification of double-stranded RNA structures in the 3'UTR of SMD targets that were similar to the STAU1 binding site (SBS) of ARF1 mRNA [67]. Two well-established SMD targets, plasminogen activator inhibitor 1 (SERPINE1) and FLJ21870 (or, ANKRD57), contain only a single Alu element in their 3'UTRs. In parallel, they also screened for lncRNAs that have a single Alu element in their sequences. They concentrated their studies on the lncRNA AF08799 that is derived from chromosome 11 of humans and potentially binds to the 3'UTRs of SERPINE1 and FLJ21870. This lncRNA was referred to as *l/2-sbsRNA1* and was found to be unaffected in expression in the STAU1-depleted conditions and was also found to be not processed by the Dicer or Drosha machineries. Interestingly, not only STAU1 depletion but knockdown of *l/2-sbsRNA1* caused an upregulation in the levels of both SERPINE1 and FLJ21870 mRNAs. Luciferase reporter assays in which the luciferase gene harbored the 3'UTRs of either of the mRNAs showed an increase in activity when *l/2-sbs RNA1* was knocked down in cells in comparison to no UTR luciferase vector control. Co-immunoprecipitation experiments in lysates of HeLa cells revealed that *l/2-sbs RNA1* interacts directly with STAU1 and SERPINE1 or FLJ21870 as well as UPF1, a protein involved in SMD. Furthermore, depletion of STAU1 reduced the interaction of *l/2-sbs RNA1* with SERPINE1 or FLJ21870 mRNAs thereby proving that binding of *l/2-sbs RNA1* to these mRNAs generates a STAU1 binding site and the binding of STUA1 stabilizes the duplex for SMD.

Further characterization of lncRNAs involved in SMD by the same group led to the identification of three other lncRNAs such as *lncRNA_BC058830* (*l/2-sbsRNA2*), *lncRNA_AF075069* (*l/2-sbsRNA3*), and *lncRNA_BC009800* (*l/2-sbsRNA4*). *l/2-sbs RNA2* targeted CDCP1 mRNA, while *l/2-sbs RNA3* and *l/2-sbs RNA4* were found to target the 3'UTR of MTAP mRNA. Knockdown of these lncRNAs individually resulted in an expected upregulation of their target mRNAs as did STAU1 or UPF1 downregulation as well. Interestingly, none of these three lncRNAs targeted SERPINE1 mRNA, revealing their specificity of action. This study was a novel one aimed at understanding the role of lncRNAs in SMD.

2.9 Biology of Natural Antisense Transcripts

Natural antisense transcripts (NATs), as the name suggests, are endogenously occurring transcripts that are coded from the opposite or antisense strand to the host gene locus. The host gene can itself be a protein-coding or a noncoding one. NATs were first discovered in prokaryotes, are classified under lncRNAs because of their length,

and are associated with ~7–30% of all genes in eukaryotes (reviewed in [68]). NATs can be either cis-NATs wherein they are encoded from the opposite strand of a gene and are complementary to the sense transcript of the same gene or trans-NATs wherein they are transcribed from a locus and show partial or full complementary to the transcript from a different locus on the same chromosome or different chromosome. Schematic representation of lncRNAs acting as NATs is depicted in Fig. 2.1d.

TSIX, the antisense transcript of X-inactive specific transcript (*XIST*), is a well-investigated NAT. To give effect to dosage compensation between males and females (females have an extra X chromosome as compared to males), *XIST* acts to coat the future inactive X chromosome in females, the inactivation in all embryonic tissues being random [69]. In order to prevent *XIST* from inactivating both the chromosomes, the future active X chromosome transcribes *TSIX*, the antisense transcript of *XIST*, thereby precluding *XIST* from acting on this chromosome. Studies performed by Zhao et al. revealed that within *Xist*, another short repeat RNA termed as *RepA*, is generated. *RepA* targets the polycomb repressive complex (PRC2) to the future inactive X chromosome and causes the establishment of repressive histone marks thereby leading to its inactivation [12]. *TSIX* acts to inhibit this interaction on the future active chromosome thus providing effect to dosage compensation. In the studies by Ohhata et al., transgenic mice containing a multiple polyadenylation signal at exon 4 of *Tsix* gene were developed; these mice prematurely terminate the transcription of *Tsix* [70]. The premature transcription termination of *Tsix* led to inappropriate activation of *Xist* on the allele which would otherwise have undergone silencing, and this activation was achieved by the establishment of active histone marks and loss of DNA methylation at the *Xist* promoter. The *Xist-Tsix* sense-antisense pair thus presents forth a wonderful paradigm for understanding the mode of action cis-NATs.

Faghihi et al. identified the antisense counterpart of β -secretase 1 (*BACE1*), an enzyme central to Alzheimer's disease pathophysiology, and termed it as *BACE1-AS* [71]. A 2 kb-long transcript, it is encoded from the strand opposite to that of *BACE1*, and it acts to regulate the levels of *BACE1* in a concordant manner. This implies that *BACE1-AS* positively co-regulates the activity of *BACE1* which was evident from the observation that si*BACE1-AS* cells showed a downregulation of both *BACE1-AS* (as expected) along with *BACE1* itself. The effect on *BACE1* was indeed observed at both the mRNA and the protein levels, and it depended on the concentration of si*BACE1-AS*, implying that higher siRNA concentrations led to higher downregulation of the *BACE1* transcript. Interestingly, a similar effect was observed in vivo when mouse brain was subjected to a continuous infusion of si*BACE1-AS*. Furthermore, in vitro experiments, when cells were exposed to various types of cell stressors like high temperature, serum starvation, hydrogen peroxide, etc. (which are known to induce AD pathology), it was observed that *BACE1-AS* levels were upregulated anywhere between 30 and 130% with a concomitant increase in *BACE1* levels as well. In samples from patients undergoing AD, *BACE1-AS* was found to have significantly higher expression as was *BACE1*, supporting the fact that *BACE1-AS* forms RNA duplex structure with *BACE1* and acts to stabilize the mRNA. Such action of a NAT to stabilize its host mRNA in AD has important

implications because not only does it enhance the activity of BACE1 in AD patients thereby playing a role in AD pathology, but also it can serve as a biomarker to detect AD in patients.

Nkx2.2AS is yet another example of a NAT that exerts a positive effect on the levels of its corresponding sense transcript *Nkx2.2* [72]. *Nkx2.2* drives the differentiation of neural stem cells toward the oligodendrocyte lineage (oligodendrocytes are a type of glial cells of the brain). Interestingly, when neural stem cell cultures were induced to differentiate by removal of Fgf-b, overexpression of *Nkx2.2 AS* led to the formation of a larger population of oligodendrocytes. It was also observed that *Nkx2.2 AS* overexpression in neural stem cells led to a significant increase in *Nkx2.2* levels thereby proving that *Nkx2.2 AS* acts to stabilize and positively regulate *Nkx2.2*.

Brain-derived neurotrophic factor (*BDNF*) is a neurotrophin required for neuronal growth, maturation, plasticity, axonal, and dendritic differentiation processes as well as in learning and memory. The antisense transcript of *BDNF* is referred to as *BDNF-AS* and is encoded 200 kb downstream of *Bdnf* locus [73]. Knockdown of *BDNF-AS* by siRNA treatment increased the mRNA and protein levels of *BDNF* although it did not affect the mRNA stability per se. When neurospheres (in vitro cultures of neural stem cells) were subjected to si*BDNF-AS*, it was observed that endogenous *BDNF* levels increased along with higher neuronal differentiation and neurite outgrowth. These results were also corroborated with in vivo studies on mouse brain wherein antagoNAT (antagonist to NAT, in this case *BDNF-AS*) indeed caused increased cellular proliferation and upregulation of *BDNF* levels. Mechanistically, *BDNF-AS* acts to induce the establishment of repressive chromatin marks on the *BDNF* locus through the recruitment of EZH2 (a component of PRC2), exemplifying the action of regulation by a NAT through epigenetic modifications.

Genome-wide analysis by tiling arrays in *Arabidopsis* revealed the presence of around 37,000 NATs, and it was observed that almost 70% of the protein-coding genes are associated with potential NATs [74]. A custom synthesized array was designed to profile the expression of NATs in various organs of *Arabidopsis* like roots, leaves, and inflorescence which led to the identification of ~15,000 NAT pairs (sense-antisense transcript pairs) in all of the organs combined with some showing tissue-specific expression as well. In order to further understand the tissue-specific regulation of NATs in *Arabidopsis*, etiolated seedlings (seedlings grown in lack of light conditions) and de-etiolated seedlings (seedlings grown in continuous white light for 1 and 6 h) were subjected to expression analysis on arrays. Interestingly, many of the candidate NATs were regulated in expression by light. SPA1 which encodes for a light signaling repressor protein was upregulated after 1 h of light exposure to seedlings along with its concordant antisense transcript, *AT2TU076050*. Intriguingly, the upregulation was seen only in the cotyledons of the seedlings, establishing the organ-specific regulation of NAT expression. Again, the mRNA coding for the protein AT3G49970, a phototropic-responsive protein, was down-regulated in hypocotyls, but its associated NAT, *AT3TU075200*, was seen to be upregulated. Such sense-antisense pairs in *Arabidopsis* were observed to be epigenetically regulated in response to light. NATs upregulated under light conditions

displayed enrichment for active histone marks; the same was true for NATs that were induced under dark conditions. Based on that, it was proposed that light-induced stimuli in *Arabidopsis* are brought about either by changes in histone modifications on the loci of NATs themselves or by NAT-guided epigenetic changes on their corresponding sense loci.

2.10 lncRNA-Mediated Regulation of Protein Activity

lncRNAs are also involved in regulation of activity of proteins involved in processes other than transcription. Hellwig and Bass (2008) reported the role of *C. elegans* lncRNA *mcs-1* in the processing of small RNAs through binding and subsequent inhibition of the Dicer enzyme [75]. lncRNA *mcs-1* is 800 nt long with its expression restricted to the hypodermis and intestine. It is transcriptionally regulated in response to food supply. *mcs-1* per se is not a substrate for Dicer due to the presence of highly branched structures at its termini, but rather *mcs-1* RNA competitively inhibits Dicer-mediated cleavage of dsRNA, as mRNA levels of several Dicer-regulated genes vary with the changes in *mcs-1* expression. An lncRNA termed as sfRNA (subgenomic flavivirus RNA) is produced by flaviviruses, such as West Nile virus, and is essential for the pathogenicity of the virus [76]. This lncRNA is a highly structured RNA of 0.3–0.5 kb, derived from the 3' untranslated region of the viral genome as a product of incomplete degradation of viral genomic RNA by cellular ribonucleases. The presence of highly conserved RNA structures at the start of the 3' untranslated region render this RNA resistant to nucleases. Mechanism of sfRNA function involves inhibition of the host XRN1 enzyme (5' to 3' exoribonuclease) during the viral RNA genome degradation and is essential for virus-induced cytopathicity and pathogenicity. Very recently, two lncRNAs have been shown to execute their biological function by the regulation of protein activity. Marchese FP et al. demonstrate the role of lncRNA *CONCR* (cohesion regulator noncoding RNA) in the regulation of sister chromatid cohesion [77]. *CONCR* is an MYC-activated, cell cycle-regulated lncRNA required for DNA replication and cell cycle progression. *CONCR* interacts with and regulates the activity of DDX11, a DNA-dependent ATPase and helicase, thereby affecting DNA replication and sister chromatid cohesion. Liu X et al. report the regulation of kinase signaling by lncRNA in the context of metabolic stress response. lncRNA neighbor of BRCA1 gene 2 (NBR2) is induced by the LKB1-AMPK pathway in the conditions of energy stress and also shown to interact with AMPK [78]. Interaction of lncRNA NBR2 with AMPK promotes its kinase activity during energy stress. Loss of NBR2 attenuates AMPK activation, leading to defective cell cycling, altered apoptosis/autophagy response, and increased tumor development. The hypoxia-regulated lncRNA linc-p21 was shown to bind HIF-1 α and VHL and disrupt the VHL-HIF-1 α interaction which leads to increased accumulation of HIF-1 α [79]. This positive feedback loop between HIF-1 α and lincRNA-p21 promotes glycolysis under hypoxia, indicating the importance of lincRNA-p21 in the regulation of the Warburg effect in tumor cells.

In addition to the modulation of gene expression through effects on mRNAs, some of the lncRNAs also regulate the activity of transcription factors. For example, lncRNA Evf2 which is transcribed from the Dlx-5/6 ultraconserved region as an alternatively spliced form of Evf-1 RNA. lncRNA Evf2 forms a stable complex with the Dlx2 transcription factor and thereby enhances the transcriptional activation of the Dlx-5/6 enhancer mediated by Dlx2 [80]. A different dimension to the regulation of protein activity by lncRNA emerges from the sequestration of transcription factors by lncRNA Gas5 which is induced under the conditions of nutrient deprivation and cellular stress [81]. Gas5 binds to the DNA binding domain of glucocorticoid receptor (GR) and acts as a decoy by competing with GR-responsive elements (GREs) in gene promoters for binding to the DNA binding domain of the GR. Therefore, Gas5 lncRNA modulates the transcriptional activity of the GR and functions as a riborepressor of the GR. Another example is the Lethe lncRNA which is regulated by TNF- α . Lethe RNA interacts with NFkB effector subunit RelA in an inducible fashion and inhibits RelA from binding to the target gene promoter [82]. The NRON (non-coding repressor of NFAT) lncRNA binds the transport receptor importin- β , and knockdown of NRON leads to nuclear accumulation of the transcription factor NFAT (nuclear factor of activated T cells) [83]. Thus NRON lncRNA competes for importin- β binding with NFAT and thus indirectly represses transcription by inhibiting the nucleocytoplasmic shuttling of NFAT.

2.11 Nuclear Architecture and Long Noncoding RNAs

The eukaryotic nucleus is a highly compartmentalized organelle with a complex dynamic organization. In order to retain the genetic material within very small nuclear volume, the cell has to package the genomic DNA into chromatin. The extensive packaging has to be performed without compromising the functional activities like transcription, DNA replication, and repair. In other words a high degree of genomic plasticity has to be maintained for efficient readout, processing, maintenance, and transfer of the genetic information which is essential for the cell to adopt different functional states. The highly condensed genomic DNA in the chromatin is associated with histone as well as nonhistone proteins. Other than regulating the function of proteins, the role of RNA in the organization of chromatin is also being investigated. In 1965, Bonner and Huang were the first to report the association of RNA with chromatin [84]. Using pea bud chromatin, they demonstrated the association of small RNA approximately 40 nt in length with the native nucleohistone. In further studies spanning three decades, various architectural functions were proposed for nuclear RNAs. These included chromatin-associated RNAs as a structural component of heterochromatin [85], role in eukaryotic chromosome structure [86], and RNA as a component of nuclear matrix and a putative role of RNA for the structural integrity of the nuclear matrix [87–89].

Scheme of nuclear architecture: The nuclear organization reflects different domains or sub-compartments which share regulatory functions. These domains

comprise of nuclear bodies and chromatin domains (reviewed in [90]). Nuclear bodies are harbored in the interchromatin space. Many of the nuclear bodies were characterized till date, including paraspeckles, cajal bodies, nucleoli, and polycomb bodies. These nuclear bodies are favorable sites for efficient biological functions. The organization of distinct nuclear bodies is similar to the organization of different organelles within the cytoplasm, and of note, the nuclear bodies lack a well-defined membrane separating them from their surroundings. Irrespective of the absence of a membrane, the nuclear bodies are structurally intact. Various studies are now shedding light on the assembly, maintenance, and regulation of nuclear bodies and also implicated a crucial role for lncRNAs in these aspects which will be discussed in detail in this review.

The chromatin domains are subdivided into transcriptionally active and inactive chromatin regions based on their gene expression status. With recent technological advancements particularly the chromosome conformation capture (3C) and other 3C-related methods such as 4C and HiC, it is clear that chromosomes occupy distinct areas in the nucleus termed as chromosome territories [91–93]. Gene-rich chromosomal regions tend to localize to the center of the nucleus and gene-poor regions near the periphery [94]. Chromatin domains with co-regulated genes often co-localize in vivo [95]. Genes located near the interphase chromosomal attachment regions to the nuclear matrix tend to be poorly transcribed [96]. Based on these observations, it can be inferred that an intimate relationship exist between the organization of genome in the nuclear space and gene activity.

2.12 lncRNAs and Nucleolar Function

The nucleolus is the nuclear domain for the rRNA synthesis, processing, and ribosomal assembly. At the onset of mitosis, i.e., during early prophase nuclear envelope breaks down and reassembles during telophase, the final stage of mitosis. This process is dependent on the RNA polymerase I transcription and the recruitment and activation of the pre-rRNA processing machinery [97]. The nucleolus contains a scaffold of tandem ribosomal DNA (rDNA) repeats. Each repeat comprises of an rDNA enhancer/promoter which is upstream of its ribosomal RNA (rRNA) coding sequence and separated by a large intergenic spacer (IGS) region. The IGS region contains a highly repetitive DNA. Characterization of different proteins in the nucleolus has shown that only 30% of the protein components have functions related to rRNA maturation. In addition to the nucleolar changes during cell cycle, the nucleolus also undergoes dramatic reorganization in response to different types of cellular stress [98, 99] (Fig. 2.2a). During different kinds of cellular stress such as acidosis, DNA damage, ribosomal stress, and serum starvation, there is preferential retention of different proteins (VHL, MDM2, DNMT1, HSP70) to the nucleolus. Therefore, in addition to being the site of rRNA maturation, the nucleolus is vital in the cellular response to stress. The nucleolar detained proteins, interact with different IGS regions downstream of the rRNA transcriptional start site. Interestingly,

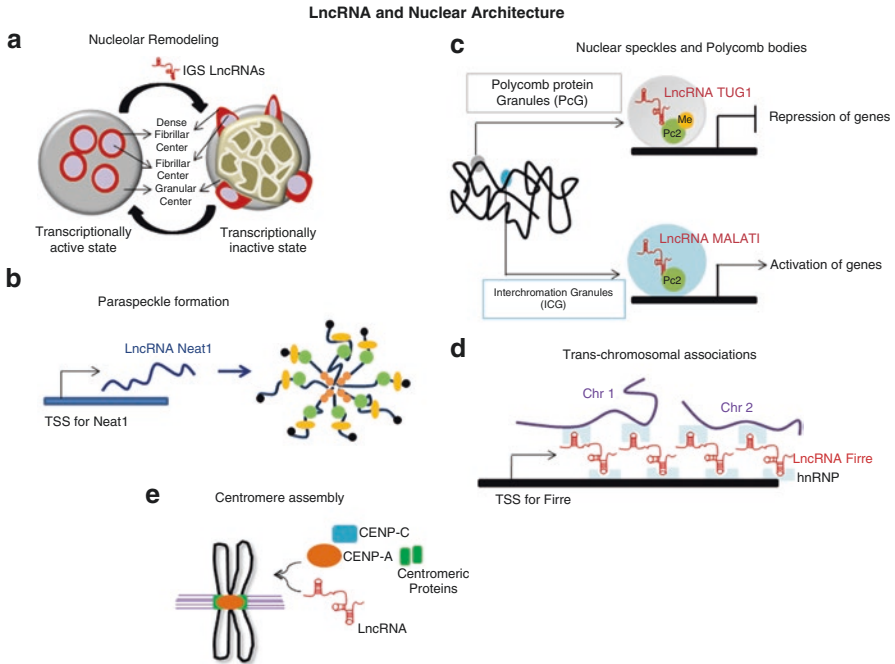


Fig. 2.2 Long noncoding RNAs and nuclear architecture. lncRNAs regulate diverse processes contributing to nuclear organization which include nucleolar remodeling (**a**), formation of paraspeckles (**b**), regulation of gene expression by association with nuclear speckles and polycomb bodies (**c**), topological organization of multichromosomal regions (**d**), and regulation of centromere assembly (**e**)

noncoding RNAs are expressed from IGS regions in response to the cellular stress that leads to the retention of the sequestered proteins at IGS loci [100]. Different IGS lncRNAs were reported to be transcribed, and each RNA responds to specific stress stimulus. Heat shock induces the transcription of IGS RNA located 16 kb (IGS₁₆RNA) and 22 kb (IGS₂₂RNA) downstream of the rDNA transcription start site. Similarly, acidosis or reduced extracellular pH induces IGS₂₈RNA located 28 kb downstream of the rDNA transcription start site. These RNAs are transcribed by the RNA Pol I machinery and processed into 300–400 nucleotide products. Expression of each IGS transcript is independent from each other with respect to the different stress stimulus. They appear to have distinct functions as the ability of one IGS RNA to sequester its target proteins was not affected by knockdown of the other IGS RNA. Strikingly, the ectopic mislocalization of these IGS RNAs in other regions of the cell also results in protein retention which highlights their function in regulating protein dynamics or mobility. Furthermore, the functional role of these IGS lncRNAs is not only limited to protein sequestration but also in the remodeling of the nucleolus (Fig. 2.2a). Using different mammalian cell lines, Jacob MD et al. describe the involvement of IGS lncRNAs in bringing about the structural and functional adaptation of the nucleolus upon heat shock or acidosis [101]. During the

latter stress conditions, the nucleolus undergoes reorganization involving de novo formation of the nucleolar detention center (DC). The DC is spatially, dynamically, and biochemically distinct sub-nucleolar structure with its architecture sustained by hydrophobic interactions among the IGS lncRNA-dependent sequestered proteins. Formation of DC leads to a reversible change in the distribution of nucleolar factors and also an arrest in the ribosome biogenesis. Formation of DC is correlated with the induction of IGS lncRNAs upon signal activation or exposure to environmental cues, and IGS RNA knockdown cells failed to form DC, and also they were unable to fully repress the nucleolar transcription upon heat shock. Therefore, the induction of IGS lncRNA by environmental cues acts as a molecular switch to regulate structure and function of the nucleolus.

Although RNA Pol I transcribes nucleolar rRNA genes, the inhibition of RNA Pol II also leads to downregulation of rRNA synthesis as well as disintegration of nucleoli [102–104]. This suggests a novel regulatory function for unknown RNA Pol II-dependent transcripts in the nucleolar organization and the Pol I transcriptional activity. In this regard, recent studies have characterized the functional role of RNA Pol II transcribed lncRNAs such as lncRNA PAPAS. The lncRNA PAPAS has been shown to be transcribed in the antisense orientation to the pre-rRNA coding region by RNA Pol II, which upon quiescence, directs the histone methyltransferase SUV4-20H2 to rRNA genes to induce histone H4 lysine 20 trimethylation (H4K20me3) and chromatin compaction [105]. The H4K20me3-mediated chromatin compaction was not just restricted to rRNA genes but was also present at other genomic elements like the IAP retrotransposons. In this case, the IAP retroelements were silenced by the IAP specific lncRNAs through the recruitment of Suv4-20 h2. In addition to quiescence, the PAPAS lncRNA also reinforces transcriptional repression of rRNA genes during hypotonic stress but by a different mechanism. During hypotonic stress, PAPAS lncRNA recruits the chromatin remodeling complex NuRD to the rDNA, leading to rDNA silencing. In another study, Herger MC et al. described the function of RNA Pol II transcripts originating from intronic Alu elements (aluRNAs) in nucleolar assembly and function. Earlier, Alu transcripts have been implicated in the regulation of gene expression and translation but not linked to the nucleolar organization [106, 107]. Through series of experiments, Herger MC et al. demonstrated that aluRNA interacts with the multifunctional nucleolar proteins nucleolin (NCL) and nucleophosmin (NPM) and tethered to chromatin and this is sufficient to target large genomic regions to the nucleolus [108]. This strongly suggests that the interaction of NCL and NPM with aluRNA is important to build up a functional nucleolus.

2.13 lncRNAs Are Critical Components of Paraspeckles

Paraspeckle is a prototypic example of mobile nuclear interchromatin sub-compartments or nuclear bodies. Paraspeckles are observed in most mammalian cultured cell lines as subnuclear granules averaging about 360 nm in diameter [109]. The

paraspeckle was first identified as a novel nuclear domain in a study to characterize the nucleolar proteome of different cultured human cells [110]. Paraspeckles are often located adjacent to splicing speckles and are marked by different proteins like PSPC1 (paraspeckle component protein 1), non-POU domain-containing octamer-binding protein (NONO, p54nrb), and SFPQ [111]. The requirement of RNA or RNA Pol II transcription in the maintenance of paraspeckle integrity was demonstrated by studies using the inhibitors like actinomycin D or D-ribofuranosylbenzimidazole (DRB) [111–113]. Several independent investigations have identified *NEAT1* lncRNA as the architectural RNA (arcRNA) of paraspeckles (Fig. 2.2b) [109–112]. Interestingly, ectopic nuclear accumulation of *NEAT1* lncRNA by tethering NEAT1 lncRNA in multiple copies to DNA leads to the de novo formation of paraspeckle [113]. In addition the role of NEAT1 lncRNA in the paraspeckle assembly was demonstrated by the direct visualization of the recruitment of the paraspeckle proteins [114]. NEAT1 has two isoforms: *NEAT1-1* (3.7 kb) and *NEAT1-2* (23 kb). In Neat1 knockout mouse embryonic fibroblasts, transient expression of Neat1–2, but not of Neat1–1, restores paraspeckle formation, demonstrating the architectural role of Neat1–2 in paraspeckle formation [115], which demonstrates a critical role for the NEAT-2 lncRNA in paraspeckle formation. However, overexpression of NEAT1–1 in cells expressing NEAT1–2 increases the number of paraspeckles, suggesting a supplementary role of NEAT1–1 in paraspeckle formation [115, 116]. Electron microscopy studies showed that the central region of NEAT1–2 is in the interior of the paraspeckle while the 5' and 3' terminus of NEAT1–2 at the paraspeckle periphery [109]. In contrast, NEAT-1 seems to localize primarily at the paraspeckle periphery. These observations collectively highlight an architectural role of NEAT1 transcripts in constraining the geometry of the paraspeckles.

2.14 lncRNAs in Nuclear Speckles and Polycomb Bodies

Nuclear speckles are nuclear domains enriched in pre-mRNA splicing factors, located in the interchromatin regions of the nucleoplasm of mammalian cells [117]. The constituents of nuclear speckles are dynamic on account of their continuous exchange with the nucleoplasm and active transcription sites. *MALAT1* lncRNA is found to be enriched in the nuclear speckles. The MALAT1 lncRNA interacts with serine/arginine (SR) splicing factors and modulates their phosphorylation and distribution to nuclear speckles, and knockdown of MALAT1 alters the splicing pattern of a subset of endogenous pre-mRNAs [118]. More importantly, the MALAT1 lncRNA has been shown to play an important role in large-scale nuclear architecture of the genome and consequently affect gene expression. This in part is achieved through lncRNA TUG1 which is localized to another distinct subnuclear body termed as polycomb group protein bodies (PcG bodies) (Fig. 2.2c). It has been proposed that activation of growth control genes occurs via inter-exchange between nuclear speckles and nuclear polycomb bodies via protein polycomb 2 (PC2) [119]. TUG1 and MALAT1 selectively interact with methylated and unmethylated Pc2 protein, respectively. Pc2 is present on growth control gene promoters. Under

conditions of reduced cell growth, TUG1 lncRNA specifically interacts with methylated Pc2. This interaction leads to the repression of growth control genes via their recruitment to PcG bodies (Fig. 2.2c). However, in the presence of growth signals, MALAT1 lncRNA interacts with unmethylated Pc2, leading to relocation of the growth control genes to the “active” environment of nuclear speckles. Hence, these lncRNAs act as key factors in the spatial regulation of specific chromatin loci.

2.15 lncRNA and Topological Organization of Multichromosomal Regions

Multiple genes which are distributed on different chromosomes can often localize within shared regions of the nucleus. These regions located in the interchromosomal nuclear domains contain genes with shared functional roles or regulated by common distant regulatory elements [120–122]. A lncRNA termed linc-RAP-1 was first identified during a loss of function screen as being important for proper adipogenesis in the mouse adipocyte precursors [123]. This lncRNA was later termed Firre (functional intergenic repeating RNA element) and shown to localize within a single nuclear domain containing many genes previously implicated in energy metabolism [124]. Firre RNA localizes across a 5 Mb nuclear domain around its site of transcription on the X chromosome (Fig. 2.2d). This single nuclear domain not only includes the Firre transcription unit but also five other chromosomal loci located on different chromosomes in trans, including chromosomes 2, 9, 15, and 17. These trans chromosomal contacts require Firre RNA since deletion of the Firre locus results in the loss of spatial proximity between Firre and its trans chromosomal binding sites. Firre is localized in a punctate fashion in the nucleus. This localization is dependent on a unique 156-bp repeating RNA domain (RRD) in the Firre sequence which is also required for a physical interaction of Firre with the nuclear matrix factor hnRNPU (Fig. 2.2d). Knockdown of hnRNPU also leads to loss of spatial proximity between Firre locus and its trans chromosomal binding sites, an effect similar to the deletion of the Firre locus itself. Random integration of Firre into different chromosomal regions leads to the emergence of new nuclear foci, suggesting that Firre may be sufficient to create a nuclear domain at the integration sites [124]. Together, these observations suggest Firre lncRNA as a nuclear organization factor as it may serve to interface with and to modulate the topological organization of multiple chromosomes.

2.16 lncRNAs and Centromere Function

Centromeres are the chromosomal regions which are the platform for the formation of kinetochore and chromosome attachment to the mitotic spindle during cell division. Irrespective of their evolutionary conserved function, the centromere identity is established epigenetically rather than being defined by the underlying DNA sequence. CENP-A (also known as CID in *Drosophila melanogaster*) is a key

epigenetic determinant of centromere identity [125]. Loss of CENP-A from cells results in the impairment of chromosome segregation, while CENP-A overexpression leads to the formation of ectopic centromeres and mislocalization of kinetochore proteins [126–128]. Deposition of CENP-A onto chromatin is carried out in a replication-independent manner unlike the canonical histones. Several studies from fission yeast to mammals; a consensus theme is emerging that lncRNA appear to play an important role in CENP-A regulation and deposition (Fig. 2.2e) [129–131]. The centromeric DNA in most organisms contains repetitive sequences or satellite sequences and form basis for characteristic pericentromeric heterochromatin [127]. Upon transcription, these repetitive and satellite sequences generate long noncoding RNAs, and these long noncoding transcripts have been shown to be involved in the initiation and maintenance of pericentromeric heterochromatin in *Drosophila*, Maize, mouse, and human cells (reviewed in [132]).

The centromeres of *D. melanogaster* contain satellite sequences which are mostly simple 5–12-bp-long repeats except for the centromere of X chromosome that contains a complex satellite repeat called as satellite III (SAT III) or 359-bp satellite and covers several megabase pairs of the acrocentric X chromosome with a 359-bp-long repeating unit [133, 134]. A recent study investigated the functional role of *D. melanogaster* SAT III region in the centromere regulation [135]. The SAT III region produces a long noncoding RNA (referred as SAT III RNA) that localizes to centromeric chromatin of the X chromosome as well as of autosomes during mitosis. SAT III RNA (~1.3 kb long) interacts with the inner kinetochore protein CENP-C and depends on CENP-C for its centromeric localization. Loss of SAT III RNA leads to mitotic defects and chromosome missegregation which is mostly attributed to the reduction of centromeric and kinetochore proteins during mitosis as well as the reduced levels of newly deposited CENP-A and CENP-C. Additionally, some SATIII RNA is also present at pericentromeres of mitotic chromosomes but not associated with chromatin. This pericentromeric SAT III RNA might also contribute to overall kinetochore structure. Hence, SAT III RNA is an integral part of centromere identity in *D. melanogaster* that influences centromere regulation epigenetically.

Centromeric repeats in maize are called CentC which are transcribed to produce transcripts that are about 900 bp long. These transcripts upon transcription associate with the maize CENP-A orthologue CENH3 and remain bound to the kinetochore and are thought to participate in the stabilization of centromeric chromatin [136]. In another recent study, Du et al. performed detailed genetic and biochemical characterization of maize inner kinetochore protein CENPC [137]. The DNA binding ability of CENPC is dependent on long single-stranded centromeric transcripts, which help in the recruitment of CENPC to the inner kinetochore.

The role of centromeric noncoding transcripts in mammalian models has been highlighted by many studies. Minor satellite repeats located on the mouse centromeres produce 4 kb-long transcripts and were implicated in regulating the centromeric function during stress response [138]. Knockdown of the transcribed murine centromeric minor satellite leads to defects in chromosome segregation [139]. These RNAs associate with CENP-A, as well as with components of the

chromosomal passenger complex (CPC): Aurora B, Survivin, and INCENP [140]. The CPC is crucial for the regulation of chromosome-microtubule attachment and the activation of the spindle assembly checkpoint. Importantly, an RNA component is necessary for the observed interaction of Aurora B with CENP-A and for the Aurora B kinase activity. Interestingly, a very recent study by Michael Blower using *Xenopus* egg extracts showed that centromeric long noncoding RNAs bind to the CPC in vitro and in vivo [141]. His work demonstrates that the centromeric lncRNAs promote normal kinetochore function by regulating the localization and activation of the CPC.

In human cells, alpha-satellite repeats have been connected to centromere function [142]. Transcript derived from the alpha-satellite repeat is 1.3 kb long and mediates the localization of CENP-C and INCENP into the nucleolus in interphase cells, and subsequently during mitosis, there is relocation of CENP-C and INCENP to centromeres (Fig. 2.2e). In a recent report, Quenet and Dalal addressed the role of a centromeric long noncoding RNA in CENP-A loading [143]. This lncRNA co-immunoprecipitates with CENP-A and its loading factor HJURP in the chromatin fraction. Downregulation of the lncRNA leads to the loss of CENP-A and HJURP at the centromeres and thus causing severe mitotic defects.

2.17 Conclusion and Outlook

From the outlined literature, it is evident that lncRNAs influence gene expression at the transcription level in cis or trans either by acting as a molecular scaffold or a decoy. lncRNA, as a molecular scaffold, has been shown to influence gene expression by targeting chromatin remodelers to specific regions across the genome. Although this phenomenon is well characterized in several biological contexts, the reasons underlying the lncRNA-mediated targeting of chromatin remodelers to specific genomic regions across the genome are very poorly understood. Likewise, lncRNAs function as decoy, involves sequestering of chromatin remodelers and transcription factors from their site of action to regulate gene expression. Although the final outcome of lncRNA actions as molecular scaffold or a decoy on gene expression is similar, it is not clear about what molecular features of lncRNA that distinguish these contrasting functions. Besides, the scaffolding and decoy functions of lncRNAs also seem to control architecture of nuclear bodies such as paraspeckles, thus influencing the global transcriptional regulation. In our view, this would be one of the most interesting aspects, needing greater insights. Similarly, lncRNAs are also considered as regulatory framework to regulate the catalytic activity of several chromatin remodelers and transcription factors through interfering with their posttranslational modifications. This raises an important question of what features of lncRNAs that impedes or promotes the catalytic activity of transcription factors and chromatin remodelers. We propose that primary sequence together with its RNA secondary structures control various aspects of protein activity, probably by blocking the catalytically active sites. Though the technologies that

address the secondary structures of RNAs have begun to be optimized, their contribution toward uncovering lncRNAs structures and their link to functions are still in their infancy. Hence, there is a greater emphasis required in optimizing technologies that can read physiological relevant secondary structures to understand their functions in various biological contexts such as molecular scaffold or decoys. Nonetheless, a decade long research on lncRNAs enabled us to know the power of dark matter of the genome in diverse biological functions. Further work should pave the way for understanding of hitherto unknown functions of lncRNAs in development and disease.

References

1. Mercer T, Mattick J (2013) Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* 20:300–307
2. Mercer T, Dinger M, Mattick J (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10:155–159
3. Geisler S, Coller J (2013) RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol* 14:699–712
4. Wang X, Song X, Glass C, Rosenfeld M (2010) The long arm of long noncoding RNAs: roles as sensors regulating gene transcriptional programs. *Cold Spring Harb Perspect Biol* 3:a003756–a003756
5. Rinn J, Chang H (2012) Genome regulation by long noncoding RNAs. *Annu Rev Biochem* 81:145–166
6. Nagano T, Fraser P (2011) No-nonsense functions for long noncoding RNAs. *Cell* 145:178–181
7. Kugel J, Goodrich J (2012) Non-coding RNAs: key regulators of mammalian transcription. *Trends Biochem Sci* 37:144–151
8. Mondal T, Rasmussen M, Pandey G, Isaksson A, Kanduri C (2010) Characterization of the RNA content of chromatin. *Genome Res* 20:899–907
9. Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, Mitra S, Mohammed A, James A, Hoberg E et al (2015) MEG3 long noncoding RNA regulates the TGF- β pathway genes through formation of RNA–DNA triplex structures. *Nat Commun* 6:7743
10. Plath K (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300:131–135
11. da Rocha S, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias N, Sanulli S, Chow J, Schulz E, Picard C et al (2014) Jarid2 is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome. *Mol Cell* 53:301–316
12. Zhao J, Sun B, Erwin J, Song J, Lee J (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322:750–756
13. Rinn J, Kertesz M, Wang J, Squazzo S, Xu X, Brugmann S, Goodnough L, Helms J, Farnham P, Segal E et al (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129:1311–1323
14. Marín-Béjar O, Marchese F, Athie A, Sánchez Y, González J, Segura V, Huang L, Moreno I, Navarro A, Monzó M et al (2013) Pint lincRNA connects the p53 pathway with epigenetic silencing by the Polycomb repressive complex 2. *Genome Biol* 14:R104
15. Klattenhoff C, Scheuermann J, Surface L, Bradley R, Fields P, Steinhauser M, Ding H, Butty V, Torrey L, Haas S et al (2013) Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* 152:570–583

16. Grote P, Herrmann B (2013) The long non-coding RNA *Fendrr* links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biol* 10:1579–1585
17. Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M, Werber M et al (2013) The tissue-specific lncRNA *Fendrr* is an essential regulator of heart and body wall development in the mouse. *Dev Cell* 24:206–214
18. Wongtrakongate P, Riddick G, Fucharoen S, Felsenfeld G (2015) Association of the long non-coding RNA steroid receptor RNA activator (SRA) with TrxG and PRC2 complexes. *PLoS Genet* 11:e1005615
19. Yang Y, Flynn R, Chen Y, Qu K, Wan B, Wang K, Lei M, Chang H (2014) Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife* 3:e02046
20. Wang K, Yang Y, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie B, Protacio A, Flynn R, Gupta R et al (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472:120–124
21. Gomez J, Wapinski O, Yang Y, Bureau J, Gopinath S, Monack D, Chang H, Brahic M, Kirkegaard K (2013) The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- γ locus. *Cell* 152:743–754
22. Hu X, Feng Y, Zhang D, Zhao S, Hu Z, Greshock J, Zhang Y, Yang L, Zhong X, Wang L et al (2014) A functional genomic approach identifies *FAL1* as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. *Cancer Cell* 26:344–357
23. Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I (2007) Characterization of a germ-line deletion, including the entire *INK4/ARF* locus, in a melanoma-neural system tumor family: identification of *ANRIL*, an antisense noncoding RNA whose expression coclusters with *ARF*. *Cancer Res* 67:3963–3969
24. Yap K, Li S, Muñoz-Cabello A, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh M, Zhou M (2010) Molecular interplay of the noncoding RNA *ANRIL* and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of *INK4a*. *Mol Cell* 38:662–674
25. Prensner J, Iyer M, Sahu A, Asangani I, Cao Q, Patel L, Vergara I, Davicioni E, Erho N, Ghadessi M et al (2013) The long noncoding RNA *SChLAP1* promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat Genet* 45:1392–1398
26. Han P, Li W, Lin C, Yang J, Shang C, Nurnberg S, Jin K, Xu W, Lin C, Lin C et al (2014) A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* 514:102–106
27. Minajigi A, Froberg J, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W et al (2015) A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349:aab2276
28. Bond A, VanGompel M, Sametsky E, Clark M, Savage J, Disterhoft J, Kohtz J (2009) Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat Neurosci* 12:1020–1027
29. Cajigas I, Leib D, Cochran J, Luo H, Swyter K, Chen S, Clark B, Thompson J, Yates J, Kingston R et al (2015) *Evf2*lncRNA/*BRG1/DLX1* interactions reveal RNA-dependent inhibition of chromatin remodeling. *Development* 142:2641–2652
30. Wang Y, He L, Du Y, Zhu P, Huang G, Luo J, Yan X, Ye B, Li C, Xia P et al (2015) The long noncoding RNA *lncTCF7* promotes self-renewal of human liver cancer stem cells through activation of Wnt signaling. *Cell Stem Cell* 16:413–425
31. Carpenter S, Aiello D, Atianand M, Ricci E, Gandhi P, Hall L, Byron M, Monks B, Henry-Bezy M, Lawrence J et al (2013) A long noncoding RNA mediates both activation and repression of immune response genes. *Science* 341:789–792
32. Hu G, Gong A, Wang Y, Ma S, Chen X, Chen J, Su C, Shibata A, Strauss-Soukup J, Drescher K et al (2016) *LincRNA-Cox2* promotes late inflammatory gene transcription in macrophages through modulating SWI/SNF-mediated chromatin remodeling. *J Immunol* 196:2799–2808
33. Tong Q, Gong A, Zhang X, Lin C, Ma S, Chen J, Hu G, Chen X (2015) *LincRNA-Cox2* modulates TNF- α -induced transcription of *IL12b* gene in intestinal epithelial cells through regulation of Mi-2/NuRD-mediated epigenetic histone modifications. *FASEB J* 30:1187–1197

34. Kim T, Hemberg M, Gray J, Costa A, Bear D, Wu J, Harmin D, Laptevich M, Barbara-Haley K, Kuersten S et al (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465:182–187
35. Ørom U, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q et al (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143:46–58
36. Koch F, Fenouil R, Gut M, Cauchy P, Albert T, Zacarias-Cabeza J, Spicuglia S, de la Chapelle A, Heidemann M, Hintermair C et al (2011) Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* 18:956–963
37. Natoli G, Andrau J (2012) Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet* 46:1–19
38. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T et al (2014) An atlas of active enhancers across human cell types and tissues. *Nature* 507:455–461
39. Antoniou M, Grosveld F (1990) Beta-globin dominant control region interacts differently with distal and proximal promoter elements. *Genes Dev* 4:1007–1013
40. Ashe H, Monks J, Wijgerde M, Fraser P, Proudfoot N (1997) Intergenic transcription and transinduction of the human beta -globin locus. *Genes Dev* 11:2494–2509
41. Melo C, Drost J, Wijchers P, van de Werken H, de Wit E, Vrieling J, Elkon R, Melo S, Léveillé N, Kalluri R et al (2013) eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol Cell* 49:524–535
42. Zhang Y, Wong C, Birnbaum R, Li G, Favaro R, Ngan C, Lim J, Tai E, Poh H, Wong E et al (2013) Chromatin connectivity maps reveal dynamic promoter–enhancer long-range associations. *Nature* 504:306–310
43. Trimarchi T, Bilal E, Ntziachristos P, Fabbri G, Dalla-Favera R, Tsiganos A, Aifantis I (2014) Genome-wide mapping and characterization of notch-regulated long noncoding RNAs in acute leukemia. *Cell* 158:593–606
44. Yang L, Lin C, Jin C, Yang J, Tanasa B, Li W, Merkurjev D, Ohgi K, Meng D, Zhang J et al (2013) lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* 500:598–602
45. Lai F, Orom U, Cesaroni M, Beringer M, Taatjes D, Blobel G, Shiekhattar R (2013) Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* 494:497–501
46. Schaukowitch K, Joo J, Liu X, Watts J, Martinez C, Kim T (2014) Enhancer RNA facilitates NELF release from immediate early genes. *Mol Cell* 56:29–42
47. Sigova A, Abraham B, Ji X, Molinie B, Hannett N, Guo Y, Jangi M, Giallourakis C, Sharp P, Young R (2015) Transcription factor trapping by RNA in gene regulatory elements. *Science* 350:978–981
48. Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H et al (2015) Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Cell Stem Cell* 16:504–516
49. Welsh I, Kwak H, Chen F, Werner M, Shopland L, Danko C, Lis J, Zhang M, Martin J, Kurpios N (2015) Chromatin architecture of the Pitx2 locus requires CTCF- and Pitx2-dependent asymmetry that mirrors embryonic gut laterality. *Cell Rep* 13:337–349
50. Kanduri C (2016) Long noncoding RNAs: lessons from genomic imprinting. *Biochim Biophys Acta* 1859:102–111
51. Wutz A, Barlow DP (1998) Imprinting of the mouse Igf2r gene depends on an intronic CpG island. *Mol Cell Endocrinol* 140:9–14
52. Sleutels F, Zwart R, Barlow D (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415:810–813
53. Mohammad F, Mondal T, Kanduri C (2009) Epigenetics of imprinted long non-coding RNAs. *Epigenetics* 4:277–286
54. Nagano T, Mitchell J, Sanz L, Pauler F, Ferguson-Smith A, Feil R, Fraser P (2008) The air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 322:1717–1720

55. Mohammad F, Mondal T, Guseva N, Pandey G, Kanduri C (2010) Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development* 137:2493–2499
56. Pandey R, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-DiNardo D, Kanduri C (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* 32:232–246
57. Kanduri C (2011) Kcnq1ot1: a chromatin regulatory RNA. *Semin Cell Dev Biol* 22:343–350
58. Mohammad F, Pandey G, Mondal T, Enroth S, Redrup L, Gyllenstein U, Kanduri C (2012) Long noncoding RNA-mediated maintenance of DNA methylation and transcriptional gene silencing. *Development* 139:2792–2803
59. Ripoche M, Kress C, Poirier F, Dandolo L (1997) Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 11:1596–1604
60. Monnier P, Martinet C, Pontis J, Stancheva I, Ait-Si-Ali S, Dandolo L (2013) H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc Natl Acad Sci* 110:20693–20698
61. Brannan C, Dees E, Ingram R, Tilghman S (1990) The product of the H19 gene may function as an RNA. *Mol Cell Biol* 10:28–36
62. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147:358–369
63. Wang L, Zhang N, Wang Z, Ai D, Cao Z, Pan H (2016) Pseudogene PTENP1 functions as a competing endogenous RNA (ceRNA) to regulate PTEN expression by sponging miR-499-5p. *Biochemistry* 81:739–747
64. Liang W, Fu W, Wang Y, Sun Y, Xu L, Wong C, Chan K, Li G, Wayne M, Zhang J (2016) H19 activates Wnt signaling and promotes osteoblast differentiation by functioning as a competing endogenous RNA. *Sci Rep* 6:20121
65. Lü M, Tang B, Zeng S, Hu C, Xie R, Wu Y, Wang S, He F, Yang S (2015) Long noncoding RNA BC032469, a novel competing endogenous RNA, upregulates hTERT expression by sponging miR-1207-5p and promotes proliferation in gastric cancer. *Oncogene* 35:3524–3534
66. Park E, Maquat L (2013) Staufen-mediated mRNA decay. *Wiley Interdiscip Rev* 4:423–435
67. Gong C, Maquat L (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 470:284–288
68. Su W, Xiong H, Fang J (2010) Natural antisense transcripts regulate gene expression in an epigenetic manner. *Biochem Biophys Res Commun* 396:177–181
69. Brown C, Hendrich B, Rupert J, Lafrenière R, Xing Y, Lawrence J, Willard H (1992) 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* 71:527–542
70. Ohhata T, Hoki Y, Sasaki H, Sado T (2007) Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. *Development* 135:227–235
71. Faghihi M, Modarresi F, Khalil A, Wood D, Sahagan B, Morgan T, Finch C, St. Laurent G III, Kenny P, Wahlestedt C (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β -secretase. *Nat Med* 14:723–730
72. Tochitani S, Hayashizaki Y (2008) Nkx2.2 antisense RNA overexpression enhanced oligodendrocytic differentiation. *Biochem Biophys Res Commun* 372:691–696
73. Modarresi F, Faghihi M, Lopez-Toledano M, Fatemi R, Magistri M, Brothers S, van der Brug M, Wahlestedt C (2012) Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nat Biotechnol* 30:453–459
74. Wang H, Chung P, Liu J, Jang I, Kean M, Xu J, Chua N (2014) Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. *Genome Res* 24:444–453
75. Hellwig S, Bass B (2008) A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. *Proc Natl Acad Sci* 105:12897–12902
76. Pijlman G, Funk A, Kondratieva N, Leung J, Torres S, van der Aa L, Liu W, Palmenberg A, Shi P, Hall R et al (2008) A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. *Cell Host Microbe* 4:579–591

77. Marchese F, Grossi E, Marín-Béjar O, Bharti S, Raimondi I, González J, Martínez-Herrera D, Athie A, Amadoz A, Brosh R et al (2016) A long noncoding RNA regulates sister chromatid cohesion. *Mol Cell* 63:397–407
78. Liu X, Xiao Z, Han L, Zhang J, Lee S, Wang W, Lee H, Zhuang L, Chen J, Lin H et al (2016) LncRNA NBR2 engages a metabolic checkpoint by regulating AMPK under energy stress. *Nat Cell Biol* 18:431–442
79. Yang F, Zhang H, Mei Y, Wu M (2014) Reciprocal regulation of HIF-1 α and LincRNA-p21 modulates the Warburg effect. *Mol Cell* 53:88–100
80. Feng J (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* 20:1470–1484
81. Kino T, Hurt D, Ichijo T, Nader N, Chrousos G (2010) Noncoding RNA Gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal* 3:ra8
82. Raponi N, Qu K, Zhang J, Mikhail M, Laberge R, Chang H (2013) A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *Elife* 2:e00762
83. Willingham A (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309:1570–1573
84. Huang R, Bonner J (1965) Histone-bound RNA, a component of native nucleohistone. *Proc Natl Acad Sci U S A* 54:960–967
85. Paul I, Duerksen J (1975) Chromatin-associated RNA content of heterochromatin and euchromatin. *Mol Cell Biochem* 9:9–16
86. Pederson T, Bhoree J (1979) Evidence for a role of RNA in eukaryotic chromosome structure. Metabolically stable, small nuclear RNA species are covalently linked to chromosomal DNA in HeLa cells. *J Mol Biol* 128:451–480
87. Nickerson J, Krochmalnic G, Wan K, Penman S (1989) Chromatin architecture and nuclear RNA. *Proc Natl Acad Sci* 86:177–181
88. Berezney R (1991) The nuclear matrix: a heuristic model for investigating genomic organization and function in the cell nucleus. *J Cell Biochem* 47:109–123
89. Ma H, Siegel A, Berezney R (1999) Association of chromosome territories with the nuclear matrix. *J Cell Biol* 146:531–542
90. Caudron-Herger M, Rippe K (2012) Nuclear architecture by RNA. *Curr Opin Genet Dev* 22:179–187
91. Dekker J (2002) Capturing chromosome conformation. *Science* 295:1306–1311
92. Dostie J, Richmond T, Arnaout R, Selzer R, Lee W, Honan T, Rubio E, Krumm A, Lamb J, Nusbaum C et al (2006) Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* 16:1299–1309
93. Zhao Z, Tavoosidana G, Sjölander M, Göndör A, Mariano P, Wang S, Kanduri C, Lezcano M, Singh Sandhu K, Singh U et al (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* 38:1341–1347
94. Cremer T, Cremer M (2010) Chromosome territories. *Cold Spring Harb Perspect Biol* 2:a003889–a003889
95. Pombo A, Branco M (2007) Functional organisation of the genome during interphase. *Curr Opin Genet Dev* 17:451–455
96. Towbin B, Meister P, Gasser S (2009) The nuclear envelope—a scaffold for silencing? *Curr Opin Genet Dev* 19:180–186
97. Hernandez-Verdun D (2011) Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus* 2:189–194
98. Boisvert F, Lam Y, Lamond D, Lamond A (2009) A quantitative proteomics analysis of sub-cellular proteome localization and changes induced by DNA damage. *Mol Cell Proteomics* 9:457–470
99. Boulon S, Westman B, Hutten S, Boisvert F, Lamond A (2010) The nucleolus under stress. *Mol Cell* 40:216–227

100. Audas T, Jacob M, Lee S (2012) Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol Cell* 45:147–157
101. Jacob M, Audas T, Uniacke J, Trinkle-Mulcahy L, Lee S (2013) Environmental cues induce a long noncoding RNA-dependent remodeling of the nucleolus. *Mol Biol Cell* 24:2943–2953
102. Haaf T, Ward D (1996) Inhibition of RNA polymerase II transcription causes chromatin decondensation, loss of nucleolar structure, and dispersion of chromosomal domains. *Exp Cell Res* 224:163–173
103. Sirri V, Hernandez-Verdun D, Roussel P (2002) Cyclin-dependent kinases govern formation and maintenance of the nucleolus. *J Cell Biol* 156:969–981
104. Burger K, Muhl B, Harasim T, Rohrmoser M, Malamoussi A, Orban M, Kellner M, Gruber-Eber A, Kremmer E, Holzel M et al (2010) Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *J Biol Chem* 285:12416–12425
105. Bierhoff H, Dammert M, Brocks D, Dambacher S, Schotta G, Grummt I (2014) Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol Cell* 54:675–682
106. Capshaw C, Dusenbury K, Hundley H (2012) Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing. *Nucleic Acids Res* 40:8637–8645
107. Singer S, Männel D, Hehlhans T, Brosius J, Schmitz J (2004) From “junk” to gene: curriculum vitae of a primate receptor isoform gene. *J Mol Biol* 341:883–886
108. Caudron-Herger M, Pankert T, Seiler J, Nemeth A, Voit R, Grummt I, Rippe K (2015) Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J* 34:2758–2774
109. 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
110. Fox A, Lam Y, Leung A, Lyon C, Andersen J, Mann M, Lamond A (2002) Paraspeckles. *Curr Biol* 12:13–25
111. Fox A (2005) P54nrb forms a heterodimer with PSP1 that localizes to Paraspeckles in an RNA-dependent manner. *Mol Biol Cell* 16:5304–5315
112. Sasaki Y, Ideue T, Sano M, Mituyama T, Hirose T (2009) MEN/noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci U S A* 106:2525–2530
113. Sunwoo H, Dinger M, Wilusz J, Amaral P, Mattick J, Spector D (2008) MEN/nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* 19:347–359
114. Mao Y, Sunwoo H, Zhang B, Spector D (2010) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* 13:95–101
115. Naganuma T, Nakagawa S, Tanigawa A, Sasaki Y, Goshima N, Hirose T (2012) Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* 31:4020–4034
116. Clemson C, Hutchinson J, Sara S, Ensminger A, Fox A, Chess A, Lawrence J (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of Paraspeckles. *Mol Cell* 33:717–726
117. Spector D, Lamond A (2010) Nuclear speckles. *Cold Spring Harb Perspect Biol* 3:a000646–a000646
118. Tripathi V, Ellis J, Shen Z, Song D, Pan Q, Watt A, Freier S, Bennett C, Sharma A, Bubulya P et al (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* 39:925–938
119. Yang L, Lin C, Liu W, Zhang J, Ohgi K, Grinstein J, Dorrestein P, Rosenfeld M (2011) ncRNA- and Pc2 Methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* 147:773–788
120. Williams A, Spilianakis C, Flavell R (2010) Interchromosomal association and gene regulation in trans. *Trends Genet* 26:188–197
121. Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* 8:104–115

122. Spilianakis C, Lalioti M, Town T, Lee G, Flavell R (2005) Interchromosomal associations between alternatively expressed loci. *Nature* 435:637–645
123. Sun L, Goff L, Trapnell C, Alexander R, Lo K, Hacisuleyman E, Sauvageau M, Tazon-Vega B, Kelley D, Hendrickson D et al (2013) Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci U S A* 110:3387–3392
124. Hacisuleyman E, Goff L, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson D, Sauvageau M, Kelley D et al (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol* 21:198–206
125. Allshire R, Karpen G (2008) Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* 9:923–937
126. Ahmad K, Henikoff S (2002) Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99:16477–16484
127. Carroll C, Straight A (2006) Centromere formation: from epigenetics to self-assembly. *Trends Cell Biol* 16:70–78
128. Mendiburo M, Padeken J, Fulop S, Schepers A, Heun P (2011) *Drosophila* CENH3 is sufficient for centromere formation. *Science* 334:686–690
129. Bouzinba-Segard H, Guais A, Francastel C (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc Natl Acad Sci* 103:8709–8714
130. Chen E, Saitoh S, Yanagida M, Takahashi K (2003) A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast. *Mol Cell* 11:175–187
131. Ohkuni K, Kitagawa K (2011) Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr Biol* 21:1695–1703
132. Rošić S, Erhardt S (2016) No longer a nuisance: long non-coding RNAs join CENP-A in epigenetic centromere regulation. *Cell Mol Life Sci* 73:1387–1398
133. Lohe A, Hilliker A, Roberts P (1993) Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Trends Genet* 9:379
134. Sun X (2003) Sequence analysis of a functional *Drosophila* centromere. *Genome Res* 13:182–194
135. Rošić S, Köhler F, Erhardt S (2014) Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol* 207:335–349
136. Topp C, Zhong C, Dawe R (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc Natl Acad Sci* 101:15986–15991
137. Du Y, Topp C, Dawe R (2010) DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet* 6:e1000835
138. Bouzinba-Segard H, Guais A, Francastel C (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc Natl Acad Sci U S A* 103:8709–8714
139. Ideue T, Cho Y, Nishimura K, Tani T (2014) Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* 19:528–538
140. Ferri F, Bouzinba-Segard H, Velasco G, Hube F, Francastel C (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res* 37:5071–5080
141. Blower M (2016) Centromeric transcription regulates Aurora-B localization and activation. *Cell Rep* 15:1624–1633
142. Wong L, Brettingham-Moore K, Chan L, Quach J, Anderson M, Northrop E, Hannan R, Saffery R, Shaw M, Williams E et al (2007) Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res* 17:1146–1160
143. Quénet D, Dalal Y (2014) A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *Elife* 3:e03254

Long Non Coding RNA Biology

Rao, M.R.S. (Ed.)

2017, XVI, 323 p. 34 illus., 32 illus. in color., Hardcover

ISBN: 978-981-10-5202-6