Review

CellPress

Special Issue: Chromatin Dynamics

Long noncoding RNAs: an emerging link between gene regulation and nuclear organization

Sofia Quinodoz and Mitchell Guttman

Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Mammalian genomes encode thousands of long noncoding RNAs (lncRNAs) that play important roles in diverse biological processes. As a class, lncRNAs are generally enriched in the nucleus and, specifically, within the chromatin-associated fraction. Consistent with their localization, many lncRNAs have been implicated in the regulation of gene expression and in shaping 3D nuclear organization. In this review, we discuss the evidence that many nuclear-retained lncRNAs can interact with various chromatin regulatory proteins and recruit them to specific sites on DNA to regulate gene expression. Furthermore, we discuss the role of specific lncRNAs in shaping nuclear organization and their emerging mechanisms. Based on these examples, we propose a model that explains how lncRNAs may shape aspects of nuclear organization to regulate gene expression.

RNA and 3D nuclear organization

Although the entire genome is present within the nucleus of every cell, distinct genes need to be accessed and expressed in different cellular conditions. Accordingly, the nucleus of each cell is a highly organized arrangement of DNA, RNA, and protein that is dynamically assembled and regulated in different cellular states [1–3]. These dynamic nuclear structures are largely arranged around functionally-related roles and often occur across multiple chromosomes [2–4]. These include large nuclear bodies (i.e., nucleolus [5,6], nuclear speckle [7], and paraspeckle [8,9]), gene–gene interactions (i.e., transcription factories [10–12] and polycomb bodies [13–16]), and promoter–enhancer interactions [17]. However, the molecular components involved in establishing this dynamic organization are still largely unknown [1–3].

It has long been suspected that RNA might play a role in organizing the structure of the nucleus. Early studies of heterogeneous nuclear RNA (hnRNA) identified a large proportion of poly(A)-modified RNA that was retained in the nucleus and was of a distinct composition from messenger RNA (mRNA) and their precursors [18–20]. Many of these RNAs were found to be localized to precise regions of the nucleus, including nuclear speckles [21] and other chromatin-associated regions [21,22]. Subsequent studies showed that global disruption of RNA transcription, but not protein translation, led to visible rearrangements of nuclear organization [23]. These studies led to the proposal that nuclear-retained RNAs might play an important structural role in the nucleus [18,21,23].

Over the past decade, many thousands of functional lncRNAs have been identified [24–27]. Recent work has highlighted that many of these lncRNAs can play important roles in diverse biological processes [28–37]. As a class, these lncRNAs are generally enriched in the nucleus and, specifically, within the chromatin-associated fraction [27,38]. Accordingly, most work on lncRNAs have focused on their role in gene regulation and, specifically, in the recruitment of chromatin regulatory proteins to genomic DNA locations [25,39,40]. In addition to this role, several recent studies have demonstrated another important role for lncRNAs in the nucleus – that is, several lncRNAs are essential for organizing distinct nuclear structures [41–50].

While lncRNAs are likely to fall into many different classes with different mechanisms [25,39,40], in this review, we focus exclusively on nuclear-retained lncRNAs that are involved in the regulation of gene expression [25,40] and in shaping 3D nuclear organization [4,35,42–45,51]. Here, we discuss the evidence demonstrating that several lncRNAs can interact with various chromatin regulatory proteins, recruit them to specific sites on DNA, modify chromatin, and regulate gene expression. Furthermore, we discuss the role of specific lncRNAs in shaping aspects of 3D nuclear organization and the emerging mechanisms by which they perform this role. Based on these examples, we synthesize the observed data into a model that may explain how some lncRNAs can shape nuclear organization to regulate gene expression – highlighting how these two apparently distinct roles may indeed occur through a shared mechanism.

Mechanisms of lncRNA regulation of gene expression through chromatin regulation

It is becoming increasingly clear that many lncRNAs can act to affect various gene expression programs [25,40]. Initial evidence for the role of lncRNAs in gene regulation came from studies of mammalian X-chromosome inactivation

Corresponding author: Guttman, M. (mguttman@caltech.edu).
Keywords: long noncoding RNA (lncRNA); chromatin regulation; genome organization; nuclear domains.

0962-8924/
 http://dx.doi.org/10.1016/j.tcb.2014.08.009

(XCI), a process that entails silencing of an entire X-chromosome in females during development [52]. This process is orchestrated by the Xist lncRNA, which is transcribed exclusively from the inactive X-chromosome (Xi) [53–55] and coats the entire Xi [56]. Importantly, genetic deletion of Xist prevents XCI [57], and induction of Xist is sufficient to initiate XCI on the same chromosome from which it is transcribed [58,59]. This silencing capability is dependent on a discrete region of the lncRNA, the A-repeat domain, whose deletion prevents transcriptional silencing without affecting Xist localization across the X-chromosome [60].

There are numerous additional examples of lncRNAs that participate in the regulation of various genes. A classic example is the *Air* lncRNA, which is responsible for regulating the *Igfr2* gene to control genetic imprinting [61–63]. In addition, *HOX* antisense intergenic RNA (HOTAIR) affects the expression of genes in the *HoxD* cluster [64] among other genes throughout the genome [33,65]. Recently, systematic studies exploring lncRNA function have shown that a large percentage of lncRNAs in the cell affect various gene expression programs [29,30,66], including those involved in embryonic development [28–30], cardiac function [31,32], immune responses [67,68], and cancer [33–37]. Based on these gene expression studies, various regulatory strategies have been proposed for lncRNAs, including the activation [47,69] and repression [34,52,64] of genes in *cis* [47,52] and in *trans* [33,64]. However, whether lncRNAs directly or indirectly regulate these target genes remains unknown.

lncRNAs can recruit chromatin regulatory proteins to genomic DNA targets

Insights into how lncRNAs can regulate gene expression initially came from studies of Xist. Specifically, female embryos containing a deletion of a component of the Polycomb Repressive Complex 2 (PRC2), which places repressive histone modifications on chromatin [70], failed to maintain proper XCI [71]. It was subsequently shown that the PRC2 complex was recruited to the entire Xi [38] and that the timing of PRC2 recruitment tightly coincides with the induction of Xist during development [52,72,73]. Importantly, deleting a discrete region of the Xist lncRNA, the B–F-repeat domain, causes a loss of PRC2 recruitment to the Xi without impacting transcriptional silencing or Xist localization across the Xi during the induction of XCI [74]. Nonetheless, there are still many open questions about Xist-mediated PRC2 recruitment. First, whether Xist physically interacts with the PRC2 complex [75,76] or indirectly recruits PRC2 [74,77,78] is still debated [78] (Box 1). Second, how Xist silences transcription and what role PRC2 may play during the induction of XCI is unclear [78], since PRC2 recruitment does not appear to be required for transcriptional silencing on the Xi [74]. Specifically, Xist mutants that disrupt the ability to recruit PRC2 (B–F repeat mutants) can still silence transcription [74], mutants that fail to silence transcription (A-repeat mutants) can still recruit PRC2 across the Xi [74,79], and Xist can induce transcriptional silencing within cells containing a genetic deletion of PRC2 [80,81]. Despite these open questions, it is clear that Xist is required to recruit the PRC2 complex across the X-chromosome [73,74,79].

This chromatin protein recruitment model may be more general beyond Xist and has been proposed for several other lncRNAs. For example, HOTAIR is thought to physically interact with PRC2, and loss-of-function of HOTAIR leads to a reduction of the PRC2-dependent histone H3 lysine 27 trimethylation (H3K27me3) repressive modifications across the *HoxD* gene cluster [33,64], suggesting that HOTAIR recruits PRC2 to these genes and may be involved in silencing their expression. Another example is *HOXA* transcript at the distal tip (*HOTTIP*), which is thought to physically interact with the WDR5 protein, and whose

Box 1. Experimental methods to define lncRNA–protein interactions

There are several common methods for purifying lncRNA–protein complexes including protein-based and RNA-based purification methods. For a more complete discussion of these methods and their strengths and limitations, see [90].

Briefly, most lncRNA–chromatin interactions [34,38,82], including the Xist–PRC2 interaction [75,76], have been identified using 'native purification' methods, which purify RNA–protein complexes under physiological conditions. The advantage of these methods is that they preserve the native complexes present in the cell. However, these methods also have several limitations, including the potential for the identification of RNA–protein interactions that form in solution, which do not reflect interactions occurring in the cell [130,131]. Because of these issues, there has been some debate about the biological significance of interactions detected by these methods [76,78,89], including the Xist–PRC2 interaction, with some arguing that they are nonspecific [78].

One way to distinguish *in vivo* interactions from interactions that form subsequently in solution is by crosslinking RNA–protein complexes in the cell and purifying the complex under denaturing conditions [131]. Methods such as CLIP utilize UV crosslinking, which crosslinks directly interacting RNA and protein molecules, to purify complexes using high-stringency wash conditions followed by separation on a denaturing SDS-PAGE gel [132,133]. A limitation of this method is that UV will not capture interactions that occur through a complex of multiple proteins [134]. This has restricted its adoption for mapping many chromatin regulatory proteins, because the precise protein within most chromatin regulatory complexes that might directly interact with a lncRNA is unknown.

Other crosslinking methods, such as formaldehyde, which crosslinks nucleic acid–protein as well as protein–protein interactions, can eliminate the need to know the exact interacting protein while enabling purification in high stringency conditions [30,135]. Indeed, several studies have used this approach to map numerous chromatin regulatory proteins, including PRC2 and WDR5, and have identified a more specific set of interactions than previously identified by native purifications [30,84]. However, adapting this formaldehyde approach to a denaturing strategy is challenging, since a denaturing SDS-PAGE gel will no longer resolve the purified complex. Furthermore, because formaldehyde crosslinks across a larger physical distances than UV, many of the interactions identified by this method might not reflect physical interactions between a lncRNA and chromatin complex [78]. For example, this approach will also identify chromatin proteins and lncRNAs that are in close proximity within a DNA locus; such proximity will likely occur for nascent transcripts and the many activating chromatin complexes bound near their transcription locus.

In the absence of the ability to define a lncRNA–protein interaction using direct crosslinking and denaturing conditions, it is unclear how to confidently define *in vivo* physical interactions using biochemical methods. In such cases, complimentary genetic methods are essential to demonstrate the functional importance of an identified lncRNA–protein interaction.

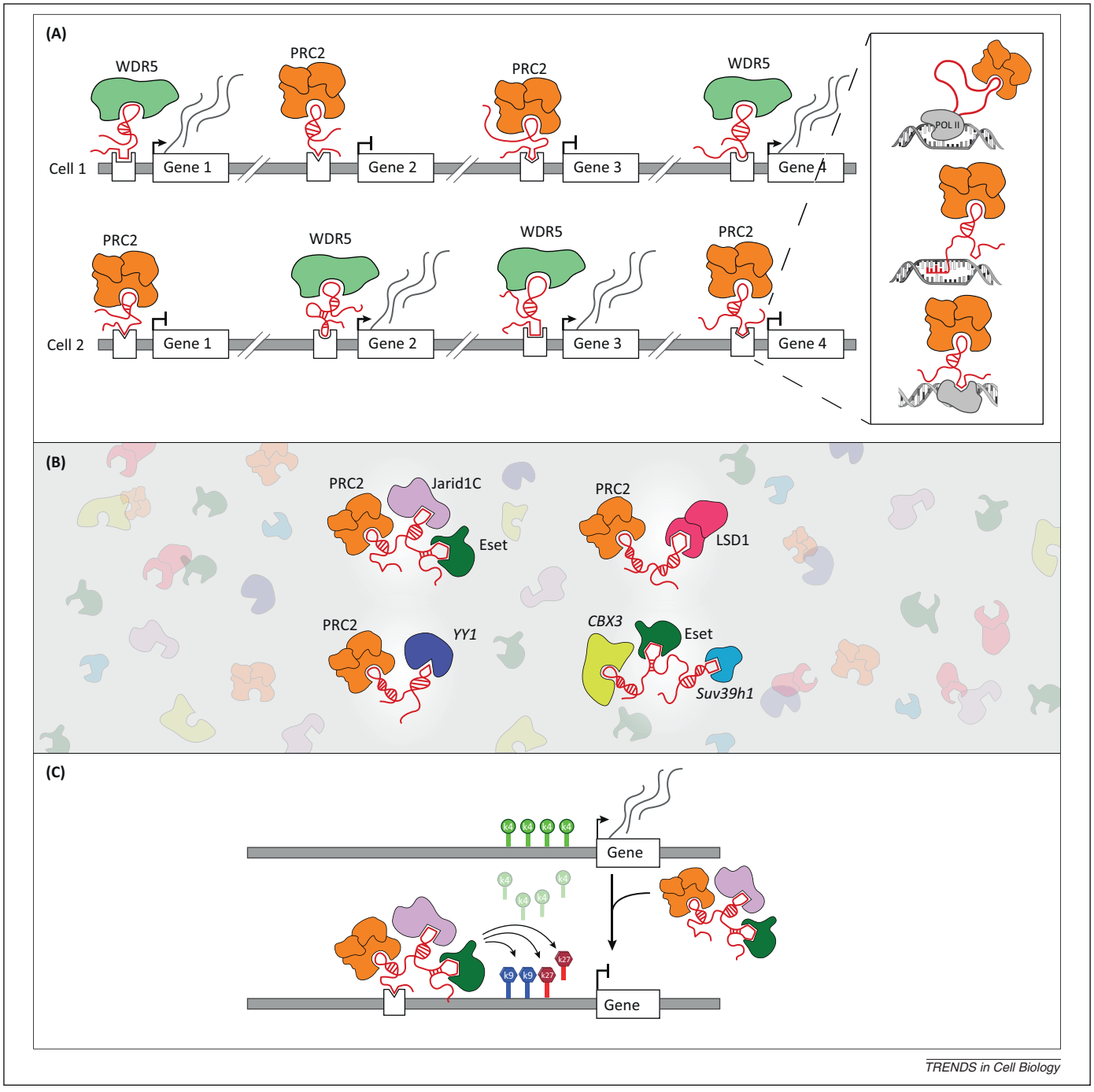

Figure 1. Long noncoding RNA (lncRNA)-mediated regulation of gene expression through the recruitment of chromatin regulatory proteins. **(A)** Different cell types express distinct lncRNAs that can differentially recruit these same chromatin regulatory proteins, including the repressive Polycomb Repressive Complex 2 (PRC2) and the activating WDR5 chromatin-modifying protein, to specific genes. Inset: lncRNAs can recruit these complexes by binding to target sites through three mechanisms: tethering to its nascent transcription locus (top panel); directly hybridizing to genomic targets (middle panel); or interacting with a DNA-binding protein (bottom panel). **(B)** Different lncRNAs can scaffold unique assemblies of chromatin regulatory complexes. lncRNAs are generally expressed at lower levels relative their associated chromatin proteins (background). **(C)** lncRNAs may act to coordinate the regulation of gene expression at specific target locations. In this illustration, a lncRNA that can scaffold PRC2, the Jarid1c histone demethylase complex, and the ESET histone methyltransferase complex may act to remove histone H3 lysine 4 trimethyl (H3K4me3) and place histone H3 lysine 27 trimethylation (H3K27me3) and histone H3 lysine 9 dimethyl (H3K9me2), thereby coordinating the repression of transcription. Abbreviations: LSD1, lysine-specific histone demethylase complex 1; YY1, Yin Yang 1; CBX3, chromobox homologue 3; PolII, RNA polymerase II.

loss-of-function leads to a reduction of its associated histone H3 lysine 4 trimethyl (H3K4me3) active histone modifications on chromatin across the *HoxA* gene cluster [47], suggesting that *HOTTIP* recruits WDR5 to these genes and may be involved in activating their expression. More generally, a large percentage of lncRNAs are thought to physically interact with various chromatin regulatory proteins, including PRC2 [30,38,82,83], WDR5 [47,84], and other readers [30,35,83,85], writers [30,35,63,86], and erasers [30,87] of chromatin modifications (Figure 1B). These examples highlight how lncRNAs may both activate and repress gene expression through a common chromatin-centric recruitment mechanism (Figure 1A).

Recently, it has been suggested that the PRC2 complex may interact with all RNAs in the cell – including lncRNAs and mRNAs [88,89]. There is considerable debate about

how many of the identified interactions between lncRNAs and chromatin proteins are specific [76,88,89] and whether these physical interactions occur through direct RNA–protein or indirect protein–protein contacts [78,90] (Box 1). However, it is increasingly clear that at least some of these lncRNA–chromatin interactions are important for lncRNA- and chromatin-mediated gene regulation. For example, mutating the RNA binding domain of WDR5 eliminates its chromatin modification and gene regulatory activities at its target sites without impacting its catalytic activity [84]. More generally, RNAi-mediated loss-of-function of several lncRNAs impacts the same genes as those impacted by loss-of-function of their associated chromatin regulatory proteins [30,38].

Together, these results suggest that many lncRNAs may recruit chromatin regulatory complexes to specific targets on genomic DNA to control gene expression (Figure 1A).

lncRNAs may scaffold multiple chromatin proteins to coordinate discrete functions

Individual lncRNAs may interact with multiple chromatin proteins simultaneously to coordinate multiple functional roles that are required to properly regulate gene expression (Figure 1B). For example, HOTAIR is thought to interact with both the PRC2 histone methyltransferase and the LSD1 histone demethylase complex [87]. This interaction may be important for coordinating the removal of activating marks (LSD1) and the addition of repressive marks (PRC2) on chromatin. More generally, more than 30 of the lncRNAs expressed in embryonic stem cells (ESCs) are thought to interact simultaneously with multiple chromatin regulatory complexes that can read, write, and erase functionally related chromatin marks [30]. Indeed, some of these ESC lncRNAs can interact with the Jarid1c histone demethylase complex, the PRC2 histone methyltransferase complex, and the ESET histone methyltransferase complex [30]. This interaction may be important for coordinating the removal of activating marks (Jarid1c) and the addition of different repressive marks on chromatin (PRC2, ESET) (Figure 1C). Importantly, many of these chromatin proteins have been shown to co-localize at specific sets of genes in ESCs, even though these proteins are not thought to directly interact with each other [91–93].

Furthermore, the Xist lncRNA is capable of coordinating at least three discrete functions to carry out its role in XCI. These functions are mediated by distinct genetic domains of the lncRNA that are required for silencing transcription (A-repeat) [60], recruitment of PRC2 (B–F-repeat) [74], and localization to chromatin (C-repeat) [94,95] – all of which are required for proper XCI. Despite these clear genetic roles, the exact molecular mechanisms by which Xist coordinates these functions remains unclear because the proteins that directly interact with Xist are still largely unknown.

Together, these results suggest that some lncRNAs may create unique assemblies of chromatin regulatory complexes and other protein complexes that do not normally form protein–protein interactions (Figure 1B). By acting as a scaffold for regulatory proteins, lncRNAs may coordinate the regulation of gene expression by recruiting a set of proteins that are required in combination for the shared regulation of a specific set of target genes (Figure 1C).

Mechanisms of lncRNA recruitment to genomic DNA

lncRNAs can recognize specific genomic DNA sites through diverse mechanisms

Three general mechanisms have been proposed for how lncRNAs that recruit protein complexes to genomic DNA can recognize specific target sites (Figure 1A). (i) RNA polymerase can tether a lncRNA to its site of transcription and, from this location, a lncRNA can act on its neighboring genes. This mechanism may explain the localization of the Neat1 lncRNA, which requires transcription to act even when large amounts of non-nascent mature RNA is present [42]. (ii) lncRNAs can interact with DNA through direct nucleic acid hybridization. This can include traditional base-pairing interactions, which explains the specificity of the telomerase RNA component for telomeric DNA repeats [96]. Additionally, lncRNAs can interact with DNA through triplex-mediated interactions, which may explain the localization of specific noncoding RNAs to ribosomal DNA promoters [97]. (iii) lncRNAs can physically interact with DNA binding proteins. Indeed the localization of *Drosophila* roX lncRNAs are dependent on their interaction with the CLAMP DNA binding protein to recognize specific DNA binding sites [65,98–100]. This mechanism may also explain the localization of Xist and Firre: both lncRNAs are thought to interact with the hnRNPU/SAF-A DNA binding protein, which is required for their localization to DNA [45,101].

However, these mechanisms – polymerase tethering, hybridization, and DNA binding protein-mediated recruitment – alone may not be sufficient to explain how a lncRNA localizes to specific sites. For example, the *roX* lncRNAs localize through their interaction with CLAMP, but they do not interact at all sites throughout the genome where CLAMP is localized [98,102]. Similarly, both Xist and Firre interact with hnRNPU/SAF-A [45,101], yet each localize to very different genomic DNA sites. This argues that specificity may not depend on a single factor, but may involve multiple independent factors, including those described above, that together provide localization specificity. Despite these examples, it remains largely unknown how most lncRNAs recognize and localize to genomic DNA.

lncRNAs can exploit the 3D conformation of the nucleus to search for targets

Recent results are pointing to a potentially general mechanism by which lncRNAs search for regulatory targets by exploiting the 3D conformation of the nucleus. For example, Xist utilizes three-dimensional nuclear organization to locate DNA target sites by first localizing to genomic sites that are in close spatial proximity to its own transcription locus [44,103]. Moving Xist to a different genomic location leads to its relocalization to new genomic target sites that are defined by their close spatial proximity to the new Xist integration site [44]. Other lncRNAs have also been shown to use spatial proximity to identify target sites [37,98,104]. *HOTTIP* localizes across the *HoxA* cluster, which is in close spatial proximity to its own transcription locus [47].

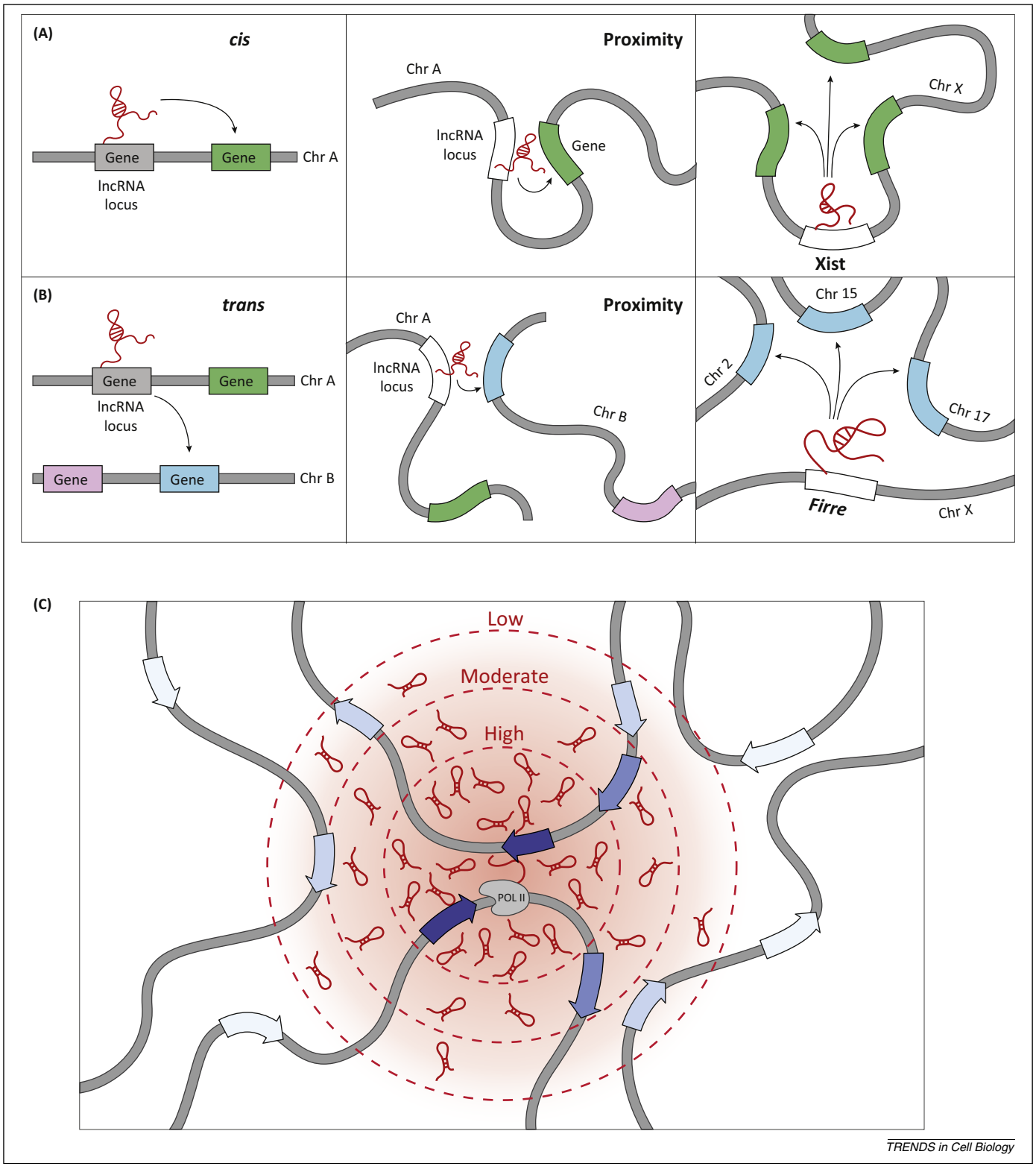

Figure 2. Long noncoding RNAs (lncRNAs) can utilize a proximity-guided search to localize to target genes. **(A)** lncRNAs can regulate genes (green box) on its own chromosome (left panel). In the nucleus, this regulation can occur if the lncRNA locus is in close physical proximity to its target sites (middle panel). For example, Xist localizes to genes across the X-chromosome (right panel). **(B)** lncRNAs can also regulate expression of genes on different chromosomes (blue box, left panel). In the nucleus, this can also occur when the lncRNA locus and its targets are in close proximity (middle panel). An example is *Firre*, which localizes to targets that are present across several chromosomes (right panel). **(C)** The concentration of a lncRNA will be highest (dark red – inner circle) near its site of transcription and will decrease (light red – outer circles) the further the distance is from its site of transcription, creating a concentration gradient of lncRNA abundance (red cloud, intensity indicates average lncRNA concentration). This spatial gradient establishes a nuclear domain with a high lncRNA concentration, where they can interact with site-specific targets (dark blue arrows). Conversely, lncRNAs outside of the nuclear domain will have a lower probability of interacting with site-specific targets (light blue arrows) due to decreased lncRNA concentration. Abbreviation: Pol II, RNA polymerase II.

This interplay between proximity-guided search and lncRNA localization is not restricted to interactions that occur on the same chromosome, but can also occur across chromosomes because regions that are present on different chromosomes can be in close spatial proximity in the nucleus [105] (Figure 2B). Indeed, the *CISTR-ACT* lncRNA localizes to sites present on the same chromosome as well as to sites on different chromosomes that are in close spatial proximity to its transcription locus [46]. This proximity-guided model may also explain the localization of HOTAIR, the first example of a *trans* regulatory lncRNA. HOTAIR is transcribed from the *HoxC* locus and regulates the expression of genes in the *HoxD* locus, which is present on a different chromosome [64]. Indeed, the *Hox* gene loci, despite being present on different chromosomes, often interact with each other in close spatial proximity within the nucleus [13,106,107]. Such a proximity-guided search model may explain the apparent observations of both *cis* and *trans* mediated regulatory mechanisms of various lncRNAs and may suggest that these apparently divergent mechanisms share a common principle of proximity within the nucleus (Figure 2).

This proximity-guided search model exploits a feature that is unique for RNA, relative to proteins, which is its ability to function immediately upon transcription. In this model, the local concentration of a lncRNA depends primarily on its spatial distance from its transcription locus, such that sites that are close will have high concentration and sites that are far will have low concentration. Yet, proximity alone is not sufficient to explain interaction, because mRNAs are also present at high concentration, but do not act, in spatial proximity to their transcription locus. Similarly, the Firre lncRNA interacts with specific DNA sites that are in spatial proximity to the Firre locus, but does not interact with all sites in spatial proximity [45]. Instead, other mechanisms, such as tethering, hybridization, or DNA binding interactions, are likely to be required for proper localization of the lncRNA to specific sites. Indeed, the *roX* lncRNAs interact with specific DNA sites, defined by the presence of CLAMP DNA elements, only when these sites are present in close spatial proximity [98,104]. These two components – proximity and sequence specificity – may explain the localization of many lncRNAs. Specifically, a lncRNA will have a high probability of interacting with a target site within a region of high concentration, but it will have a low probability of interacting with a target site within a region of low concentration – even if it has a high affinity for that site (Figure 2C). Importantly, such a strategy might explain how lncRNAs, which are generally of lower abundance, could reliably identify their target genes by searching in close spatial proximity to their transcription locus rather than searching across the entire nucleus.

lncRNAs are essential for the establishment and maintenance of nuclear domains

Recent studies have highlighted another link between lncRNAs and nuclear organization – that is, several lncRNAs have been shown to play a critical role in bringing together DNA, RNA, and proteins to actively shape some aspects of 3D nuclear organization. We discuss examples of lncRNAs that establish nuclear domains across various levels of organization from nuclear bodies to enhancer–promoter interactions below (Figure 3).

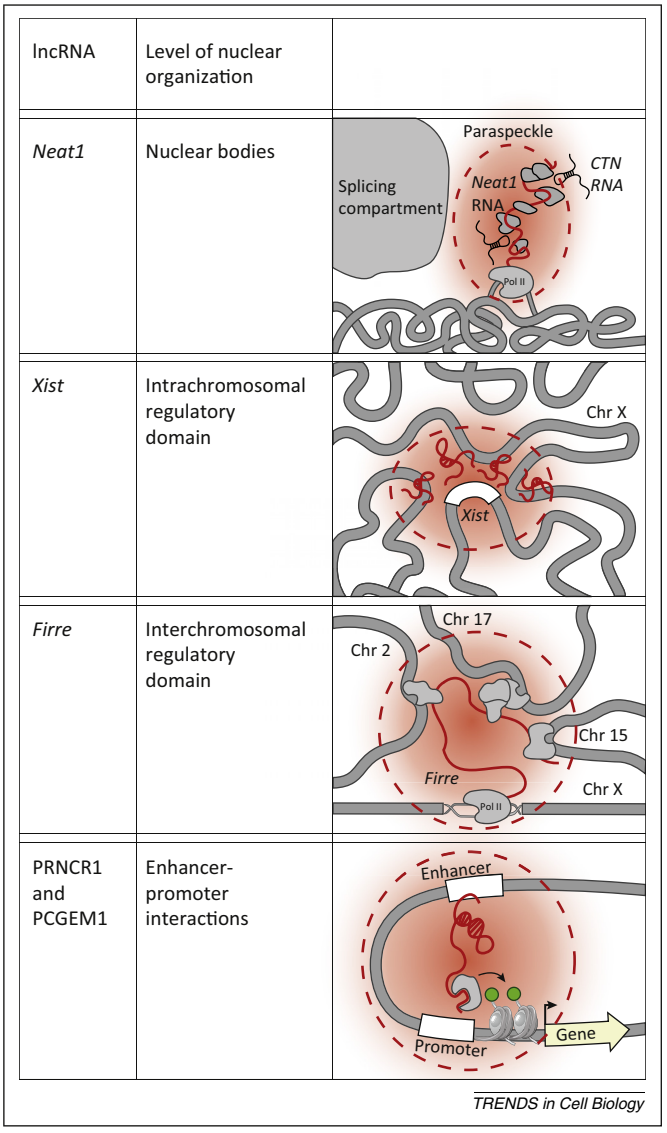

Figure 3. Long noncoding RNAs (lncRNAs) can shape 3D nuclear architecture across various levels of organization. **(A)** Actively transcribed *Neat1* (red line) is required to establish the formation of the paraspeckle nuclear body (red cloud), which is an RNA–protein (gray) nuclear domain that is the site of nuclear retention of RNAs, such as the *CTN RNA* (black). **(B)** Xist (red line) establishes an intrachromosomal nuclear domain (red cloud) by nucleating near its transcription site (white box) and spreading to DNA sites in spatial proximity to its locus. **(C)** *Firre* establishes an interchromosomal nuclear domain and brings together targets on chromosomes 2, 15, and 17 into close physical proximity to its transcriptional locus on the X-chromosome. **(D)** lncRNAs acting at enhancers, such as the PRNCR1 and PCGEM1 lncRNAs, maintain the interaction between enhancer and promoter regions and may do this by interacting with proteins that can modify chromatin.

Nuclear bodies: Neat1 establishes the paraspeckle

The paraspeckle consists of various RNAs and proteins that are spatially co-localized and is thought to be the site of nuclear retention of adenosine-to-inosine edited mRNAs [8,9]. Recent studies have demonstrated an essential role for the *Neat1* lncRNA in forming paraspeckles [42,43,108]. Specifically, Neat1 has been found to localize within the paraspeckle [43,108,109], and its loss of function leads to a loss of the paraspeckle domain [43,108]. Conversely, induction of Neat1 is sufficient to establish the paraspeckle domain

[42]. Furthermore, recruitment of Neat1 to a transgenic site is sufficient to create paraspeckles at that location [42,110]. Indeed, synthetically tethering Neat1 to a genomic DNA region is sufficient to form paraspeckles, but tethering the paraspeckle-associated proteins, such as PSP1, to DNA is not sufficient to assemble paraspeckles [110]. Neat1 transcription is required for establishing and maintaining paraspeckles by recruiting paraspeckle-associated proteins to the *Neat1* genomic locus [42]. Accordingly, disruption of Neat1 transcription, even without a reduction in overall Neat1 levels, leads to the loss of paraspeckles [42].

Together, these studies demonstrate that Neat1 plays an architectural role in the establishment and maintenance of the paraspeckle nuclear domains by seeding at its transcription locus and recruiting associated proteins to create an RNA–protein nuclear compartment.

Intrachromosomal regulatory domains: Xist compacts the X-chromosome

During XCI, the inactive X-chromosome is compacted and relocated to the periphery of the nucleus to form an intrachromosomal domain, termed the Barr body [52]. This 3D restructuring of the Xi is carried out by the Xist lncRNA [52,111]. Indeed, integrating Xist into transgenic locations, including on autosomes, is sufficient to silence, compact, and reposition the chromosome on which Xist is integrated [59,112]. Xist spreads from its transcription locus to initial sites that are in close spatial proximity [44]. From these sites, Xist then spreads across the entire X-chromosome. This spreading process is known to involve significant changes to chromosome architecture across the X-chromosome [52,111]. These structural changes depend on the A-repeat domain of Xist, the same domain required for silencing transcription, because deletion of the A-repeat leads to the exclusion of actively transcribed regions from the silenced X-chromosome territory [44,51].

Together, these studies demonstrate that Xist is necessary for restructuring genomic DNA regions to establish an RNA-mediated silenced nuclear compartment. Xist performs this function by spreading across the X-chromosome and repositioning genes into the silenced Xist compartment [44].

Interchromosomal regulatory domains: Firre forms a trans-chromosomal compartment

Multiple genes that are present on different chromosomes can often localize within shared regions of the nucleus. These interchromosomal nuclear domains are often defined by the presence of genes with shared functional roles or regulation by common factors [105,113,114]. Recently, a lncRNA termed Firre was identified based on its role in adipogenesis [66]. This lncRNA was shown to localize within a single nuclear domain containing many genes previously implicated in energy metabolism [45,66]. This single nuclear domain includes the Firre transcription locus on the X-chromosome as well as at least five genes that are located on different chromosomes including chromosomes 2, 9, 15, and 1745. Importantly, deletion of the *Firre* locus results in reduced co-localization of the *trans*-chromosomal contacts within this nuclear domain [45]. 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 compartment at its integration sites [45]. Taken together, these results suggest that Firre is required to maintain, and may even be required to establish the formation of a trans-chromosomal nuclear compartment containing target genes of shared function.

Enhancer–promoter interactions: lncRNAs acting at enhancers can promote chromosomal looping

Gene regulation involves physical interactions between distal enhancer regions and the promoters that they regulate. Recent studies have shown that some lncRNAs can act at active enhancer regions. Several lncRNAs have been proposed to play a role in mediating chromosomal interactions between an enhancer region and its associated promoter [35,37,48,115–117]. For example, estrogen-induced [116] and androgen-induced lncRNAs [35] have been shown to maintain DNA looping between enhancer and promoter regions and, through this interaction, promote gene activation of estrogen-responsive genes and androgen-receptor-activated genes, respectively. Together, these results suggest that some lncRNAs acting at enhancers are required to maintain the 3D chromosomal looping between an enhancer and its associated promoter.

While much is still unknown about how these lncRNAs acting at enhancers work, initial insights are emerging from two specific lncRNAs that are highly expressed in prostate cancer [35]. The *PRNCR1* lncRNA binds to the enhancer regions of androgen-receptor regulated genes and is thought to recruit the DOT1L histone methytrasferase to the enhancer. This chromatin protein recruitment in turn recruits a second lncRNA, *PCGEM1*, to the same region. The *PCGEM1* lncRNA is thought to interact with Pygo2, an H3K4me3 reader that can recognize methylation groups on active promoter regions [35]. Through the recruitment of these proteins, these lncRNAs appear to facilitate looping between the enhancer and promoter regions, leading to the subsequent activation of the target gene [35]. These results suggest that lncRNAs acting at enhancers may recruit chromatin regulatory proteins to create high affinity interactions between different regions of DNA and, through this, act to reposition enhancer and promoter regions into close spatial proximity.

Collectively, these results and others [41] demonstrate that several lncRNAs play an important role in establishing and maintaining higher-order nuclear structures across various levels of nuclear organization from nuclear bodies to enhancer–promoter interactions.

A proposed model for lncRNA-mediated organization of nuclear structure

While there is some evidence that lncRNAs can recruit chromatin regulators, modify chromatin structure, regulate gene expression, search in spatial proximity, and reposition genes into a nuclear domain, how these mechanisms work together to create a dynamic nuclear compartment remains unclear. Important insights can be derived from studies of nuclear body formation, which depends on many molecules – including RNA, DNA, and protein – coming together into a single nuclear region [4,118,119]. This process requires the localization of an initial

nucleating factor, which seeds organization and recruits other factors to this location [4,119]. For example, tethering individual RNAs or proteins that are present in the Cajal bodies to a random location in the genome is sufficient to seed the formation of a new Cajal body at that site [110,120]. In the context of well-studied nuclear bodies, such as the Cajal body and the nucleolus, the proteins involved have domains that allow them to self-interact, thereby creating preferential interactions between molecules of the same identity [118,119,121,122]. This self-organization creates a high local concentration of a defined set of molecules within a spatially confined region around the initial nucleating factor [119].

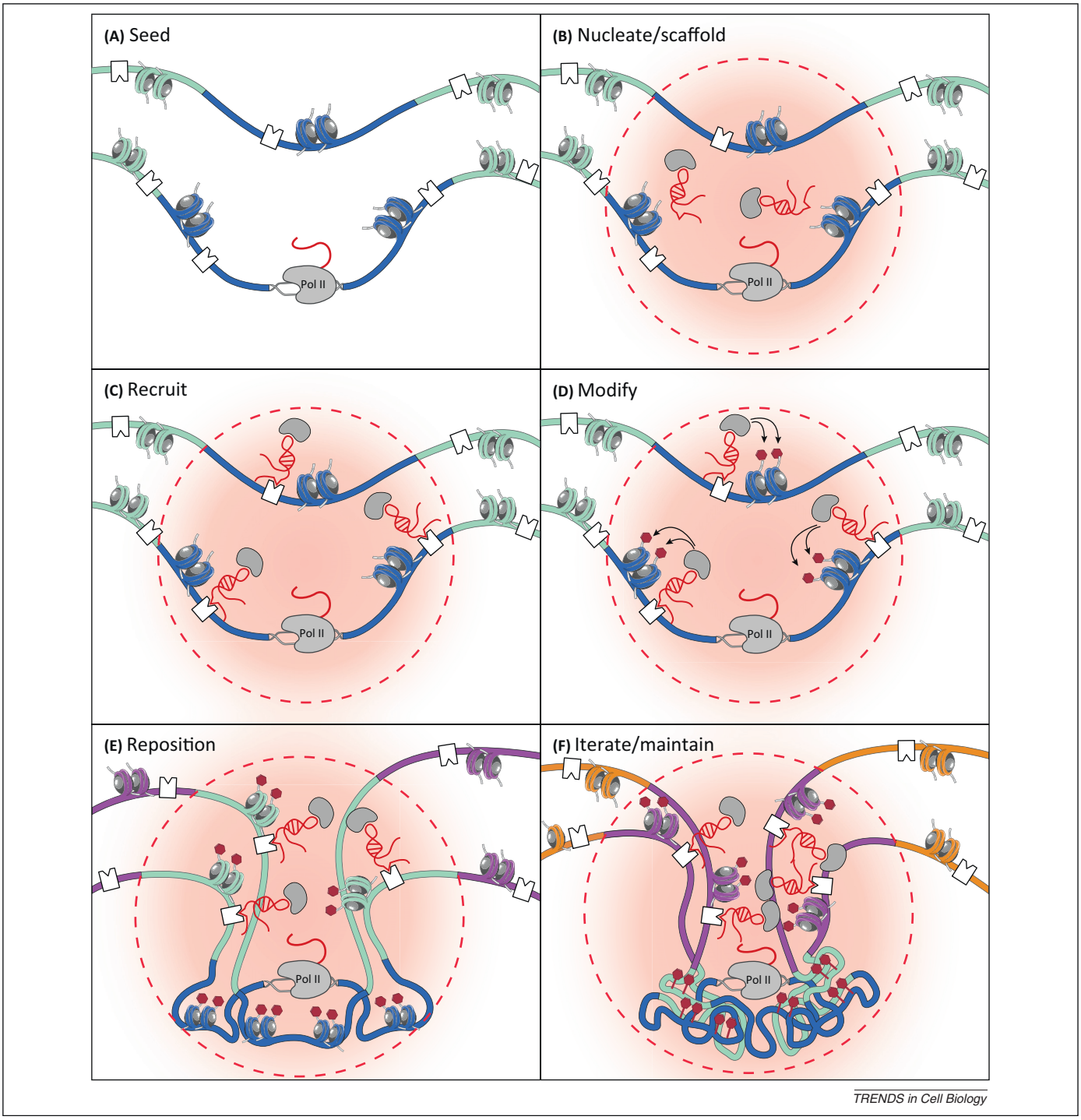

Figure 4. A model for how long noncoding RNAs (lncRNAs) can dynamically shape nuclear organization. The proposed steps involved in lncRNA mediated assembly of nuclear organization roughly based on the proposed models for the *Neat1* [42] and Xist [44] lncRNAs. **(A)** Transcription of a lncRNA can seed the formation of a lncRNA nuclear domain. **(B)** lncRNAs can bind to proteins in the nucleus (gray circles) to scaffold protein complexes. Formation of these complexes will nucleate the formation of a spatial compartment (red cloud, broken lines) near the transcriptional locus of the lncRNA. **(C)** lncRNAs can bind to specific DNA sites (white squares) to recruit lncRNA–protein complexes to target sites. **(D)** By recruiting these complexes to DNA, lncRNAs can guide chromatin modifications (blue histones), such as repressive histone modification (red marks). **(E)** Modified chromatin may be compressed and repositioned into a new nuclear region. **(F)** As the lncRNA continues to be transcribed from its transcriptional locus, it may iteratively bind to DNA sites (green regions), modify target sites, and reposition DNA into the lncRNA nuclear domain. This continuous process may act to maintain the nuclear domain established by a lncRNA. Abbreviation: PolII, RNA polymerase II.

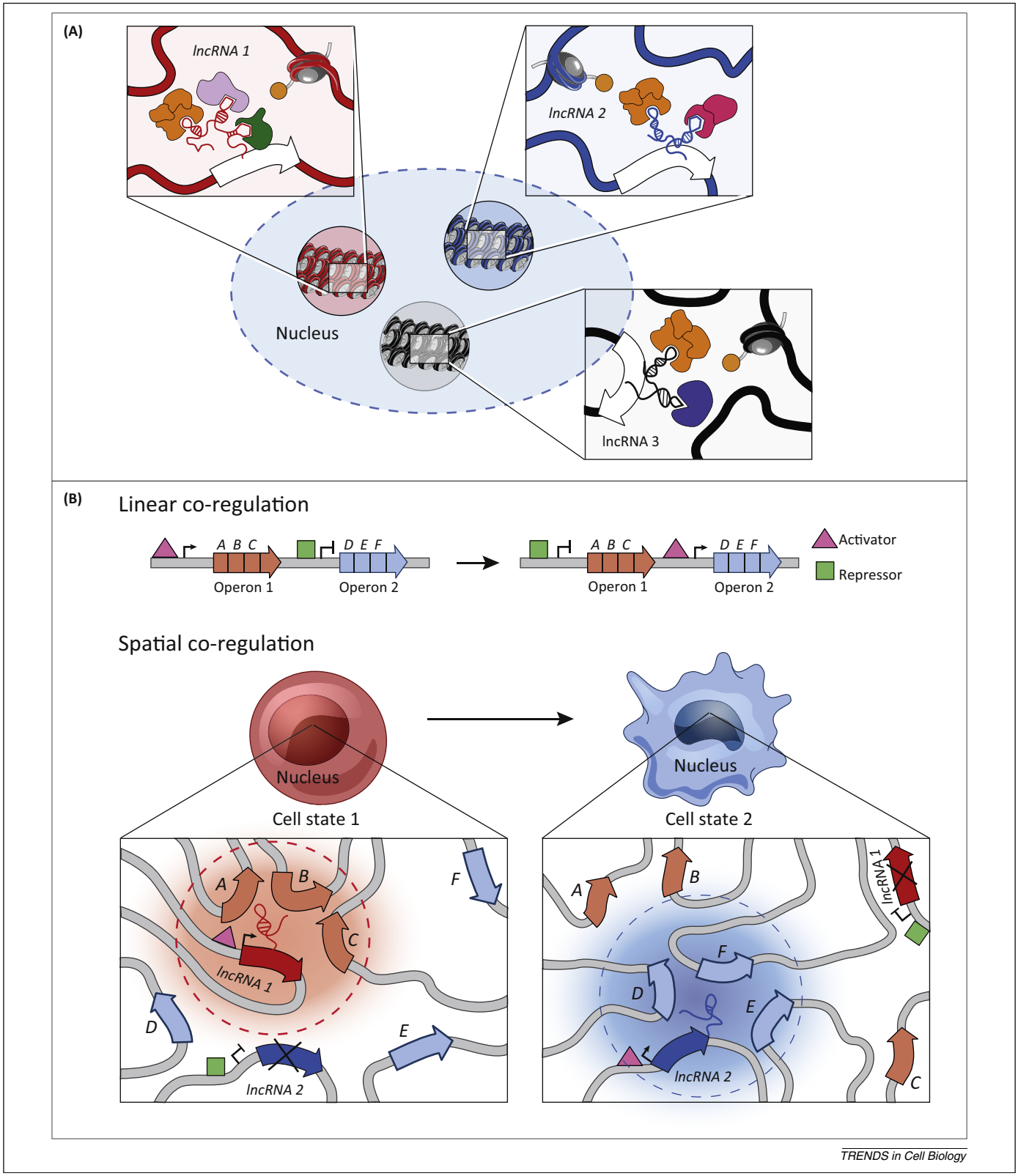

Figure 5. A hypothesis for how long noncoding RNAs (lncRNAs) may act to assemble dynamic and specific nuclear domains. **(A)** Nuclear domains that share the same proteins can interact in different regions of the nucleus. Zoom-in panels: we hypothesize that different lncRNAs may act to distinguish between these domains by scaffolding and assembling distinct domains. **(B)** Through linear co-regulation, operons can simultaneously regulate sets of genes (*A, B, C* and *D, E, F*) with shared regulatory functions. Activators (pink triangles) and repressor (green boxes) control operon expression under a particular cell state. We hypothesize that through spatial co-regulation, lncRNAs may nucleate the formation of nuclear domains to co-localize target genes upon induction of lncRNA expression. For example, upon induction of *lncRNA1*, genes *A, B*, and *C* are co-regulated in a nuclear domain (red cloud, broken lines). Under a different cell state, *lncRNA1* expression is repressed, leading to the breakdown of the *lncRNA1* nuclear domain and expression of *lncRNA2* leads to formation of another nuclear domain (blue cloud, broken lines) containing genes *D, E*, and *F*.

This process may also explain the assembly of other functional nuclear structures: DNA containing common chromatin modification patterns [16], or DNA that is bound by shared proteins, such as PRC2 [13–15] or various transcription factors [10–12], can cluster together in 3D proximity. While the exact mechanism that leads to the formation of these particular long-range interactions is largely unclear [123], it appears that a similar self-organization property may be involved because molecules of shared identity preferentially interact in 3D proximity [124,125].

Based on the studies discussed above, we propose a model for how lncRNAs may organize nuclear architecture (Figure 4). This model is an extension of those originally proposed for Neat1 [42] and Xist [44]. In this model, lncRNAs can seed organization by creating domains of high local lncRNA concentration near their site of transcription. This would allow the lncRNAs to scaffold various protein complexes and thereby nucleate a lncRNA–protein complex assembly, increasing the effective concentration of proteins within this domain [4,42]. lncRNAs can then interact with high affinity target sites to achieve specificity and recruit lncRNA–protein complexes to specific target sites. At these targets, lncRNA–protein complexes may modify the chromatin state and, through this, may act to reposition DNA sites into new nuclear domains of shared chromatin modification or protein occupancy [44]. Importantly, whether chromatin modifications or other mechanisms, such as self-organization based on the recruitment of shared protein complexes, are what drive repositioning remains to be tested. This proposed model may not be restricted to the formation of DNA compartments, but may also explain the spatial assembly of RNA and protein domains in the nucleus through a similar lncRNA-centric mechanism [42,110,122].

This process of lncRNA spreading and repositioning may involve iterative steps by which the lncRNA, while actively transcribed, can continue to seed, nucleate, modify, and reposition genes into an expanding nuclear domain [44]. For example, Xist spreads to new sites on the X-chromosome by interacting in spatial proximity with the genes that have not yet been silenced and then repositioning these genes into the growing silenced nuclear compartment [44]. Once established, the lncRNA may maintain this domain through continued transcription from a location in close spatial proximity to the newly formed compartment, similar to how Neat1 is required to maintain paraspeckles through an ongoing process of transcription [42].

Possible implications of lncRNA-mediated nuclear organization in gene regulation

It is increasingly clear that there are functional nuclear domains that contain shared chromatin modification patterns [16] or protein occupancy [2,10,12,14]. However, not all DNA that is modified or bound by a specific protein in the nucleus is spatially localized within a single nuclear domain [13,16]. We hypothesize that different lncRNAs may establish these specific nuclear domains by scaffolding and recruiting distinct combinations of proteins. (Figure 5A). For example, the nucleus contains multiple discrete functional domains that are enriched for polycomb protein occupancy (polycomb bodies) [14,15]; one such domain is the inactive X-chromosome [79], which is established by a specific lncRNA and is spatially distinct from other polycomb-enriched domains in the nucleus.

While nuclear organization is known to be highly dynamic between cell states, how this organization is dynamically established during various processes, such as cellular differentiation, remains unclear [11,106,126,127]. We hypothesize that some lncRNAs might act as 'organizational centers' to establish cell-type specific nuclear domains that organize genes of similar function in close 3D proximity. Such a role is consistent with the observation that lncRNAs exhibit extraordinary cell-type specificity [26,27,128], in contrast to proteins, which are often reused in multiple cellular contexts [129]. In this model, nuclear compartments can be dynamically organized simply through the activation or repression of a single lncRNA gene (Figure 5B).

This hypothesized role of lncRNAs as organizational centers might represent an ideal strategy for how nuclear-localized lncRNAs could act to regulate gene expression. Because lncRNAs are generally expressed at low abundance, the probability of coordinately finding multiple target genes that are distributed throughout the nucleus would be low, potentially leading to heterogeneous expression of these genes. There are two theoretical solutions: increase lncRNA abundance or cluster target genes in spatial proximity. While both approaches solve the challenge of finding distributed genes, increasing the levels of a lncRNA may not be an optimal solution because this may lead to subsaturation of a lncRNA scaffold with its associated regulatory proteins (Figure 1B). Therefore, lncRNA regulation of multiple distributed genes requires a tradeoff between the optimality of finding all genes (high lncRNA expression) with the optimality of interacting with all required regulatory proteins (low lncRNA expression). Spatial clustering would provide an ideal solution because it would enable a lncRNA to easily find all of its targets based on spatial proximity, where the lncRNA is in high local concentration, while ensuring saturation of the lncRNA regulatory complexes to coordinately regulate all of its target genes.

Concluding remarks

While the role of lncRNAs in establishing nuclear organization is attractive, many questions remain. Currently, there are only a few examples of lncRNAs that organize nuclear domains, and even for these few lncRNAs, how they organize these nuclear domains is largely unknown. Future studies will be required to determine whether this role may be a more general role for nuclear-retained lncRNAs and whether there may be general mechanistic principles by which lncRNAs act to shape nuclear domains. In particular, it will be important to identify additional lncRNA-mediated nuclear domains and characterize the dynamics of their formation across various cellular conditions. Such examples will allow us to dissect the precise mechanisms by which lncRNAs can organize nuclear domains and determine the various components required for domain assembly. To address these questions, it will be important to develop experimental systems, such as

inducible lncRNA systems that enable precisely controlled formation of the associated nuclear domain, to dissect dynamic nuclear organization at the molecular level. Such experimental systems will enable the systematic perturbation of a lncRNA, including deletion of specific protein binding regions, and the measurement of their roles in the establishment and maintenance of nuclear domains. Finally, it will be essential to determine the role that lncRNA-mediated regulation of nuclear organization plays in the control of gene expression. While much work remains to be done, it is now clear that the roles of lncRNAs in regulating gene expression and establishing nuclear organization may be more tightly linked than previously appreciated.

Acknowledgments

We thank Mario Blanco, Jesse Engreitz, Manuel Garber, Shari Grossman, Colleen McHugh, and Klara Stefflova for helpful discussions, critical reading, and suggestions as well as Howard Chang for discussions about *roX* localization, John Rinn for discussions about proximity mechanisms, Kathrin Plath for discussions about Xist and nuclear organization, and Sigrid Knemeyer for illustrations and assistance with figures. S.Q. is a Howard Hughes Medical Institute Gilliam Fellow. Work in the Guttman lab is funded by an NIH Director's Early Independence Award, the Searle Scholars program, the Edward Mallincrodt Jr Foundation, and the Sidney Kimmel Foundation.

References

1 Gibcus, J.H. and Dekker, J. (2013) The hierarchy of the 3D genome. *Mol. Cell* 49, 773–782
2 Cavalli, G. and Misteli, T. (2013) Functional implications of genome topology. *Nat. Struct. Mol. Biol.* 20, 290–299
3 Misteli, T. (2007) Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800
4 Mao, Y.S. *et al.* (2011) Biogenesis and function of nuclear bodies. *Trends Genet.* 27, 295–306
5 McStay, B. and Grummt, I. (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* 24, 131–157
6 Melese, T. and Xue, Z. (1995) The nucleolus: an organelle formed by the act of building a ribosome. *Curr. Opin. Cell Biol.* 7, 319–324
7 Lamond, A.I. and Spector, D.L. (2003) Nuclear speckles: a model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* 4, 605–612
8 Fox, A.H. and Lamond, A.I. (2010) Paraspeckles. *Cold Spring Harb. Perspect. Biol.* 2, a000687
9 Fox, A.H. *et al.* (2002) Paraspeckles: a novel nuclear domain. *Curr. Biol.* 12, 13–25
10 Schoenfelder, S. *et al.* (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* 42, 53–61
11 Wei, Z. *et al.* (2013) *Klf4* organizes long-range chromosomal interactions with the *oct4* locus in reprogramming and pluripotency. *Cell Stem Cell* 13, 36–47
12 Osborne, C.S. *et al.* (2007) *Myc* dynamically and preferentially relocates to a transcription factory occupied by *Igh*. *PLoS Biol.* 5, e192
13 Bantignies, F. *et al.* (2011) Polycomb-dependent regulatory contacts between distant *Hox* loci in *Drosophila*. *Cell* 144, 214–226
14 Cheutin, T. and Cavalli, G. (2014) Polycomb silencing: from linear chromatin domains to 3D chromosome folding. *Curr. Opin. Genet. Dev.* 25C, 30–37
15 Delest, A. *et al.* (2012) Polycomb: a paradigm for genome organization from one to three dimensions. *Curr. Opin. Cell Biol.* 24, 405–414
16 Sexton, T. *et al.* (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472
17 Smallwood, A. and Ren, B. (2013) Genome organization and long-range regulation of gene expression by enhancers. *Curr. Opin. Cell Biol.* 25, 387–394
18 Herman, R.C. *et al.* (1976) Message and non-message sequences adjacent to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. *Cell* 7, 429–437
19 Perry, R.P. *et al.* (1974) Synthesis and turnover of nuclear and cytoplasmic polyadenylic acid in mouse L cells. *J. Mol. Biol.* 82, 315–331
20 Brawerman, G. and Diez, J. (1975) Metabolism of the polyadenylate sequence of nuclear RNA and messenger RNA in mammalian cells. *Cell* 5, 271–280
21 Huang, S. *et al.* (1994) In vivo analysis of the stability and transport of nuclear poly(A)+ RNA. *J. Cell Biol.* 126, 877–899
22 He, D.C. *et al.* (1990) Core filaments of the nuclear matrix. *J. Cell Biol.* 110, 569–580
23 Nickerson, J.A. *et al.* (1989) Chromatin architecture and nuclear RNA. *Proc. Natl. Acad. Sci. U.S.A.* 86, 177–181
24 Guttman, M. *et al.* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223–227
25 Guttman, M. and Rinn, J.L. (2012) Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339–346
26 Cabili, M.N. *et al.* (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927
27 Derrien, T. *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789
28 Sauvageau, M. *et al.* (2013) Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife* 2, e01749
29 Ulitsky, I. *et al.* (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537–1550
30 Guttman, M. *et al.* (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477, 295–300
31 Klattenhoff, C.A. *et al.* (2013) Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* 152, 570–583
32 Han, P. *et al.* (2014) A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* (http://dx.doi.org/10.1038/nature13596)
33 Gupta, R.A. *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076
34 Huarte, M. *et al.* (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142, 409–419
35 Yang, L. *et al.* (2013) lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* 500, 598–602
36 Gutschner, T. *et al.* (2013) The noncoding RNA *MALAT1* is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* 73, 1180–1189
37 Trimarchi, T. *et al.* (2014) Genome-wide mapping and characterization of notch-regulated long noncoding RNAs in acute leukemia. *Cell* 158, 593–606
38 Khalil, A.M. *et al.* (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11667–11672
39 Wang, K.C. and Chang, H.Y. (2011) Molecular mechanisms of long noncoding RNAs. *Mol. Cell* 43, 904–914
40 Rinn, J.L. and Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166
41 Bergmann, J.H. and Spector, D.L. (2014) Long non-coding RNAs: modulators of nuclear structure and function. *Curr. Opin. Cell Biol.* 26, 10–18
42 Mao, Y.S. *et al.* (2011) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* 13, 95–101
43 Clemson, C.M. *et al.* (2009) An architectural role for a nuclear noncoding RNA: *NEAT1* RNA is essential for the structure of paraspeckles. *Mol. Cell* 33, 717–726
44 Engreitz, J.M. *et al.* (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341, 1237973
45 Hacisuleyman, E. *et al.* (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA *Firre*. *Nat. Struct. Mol. Biol.* 21, 198–206
46 Maass, P.G. *et al.* (2012) A misplaced lncRNA causes brachydactyly in humans. *J. Clin. Invest.* 122, 3990–4002
47 Wang, K.C. *et al.* (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124

48 Lam, M.T. *et al.* (2014) Enhancer RNAs and regulated transcriptional programs. *Trends Biochem. Sci.* 39, 170–182
49 Shichino, Y. *et al.* (2014) Meiotic long non-coding meiRNA accumulates as a dot at its genetic locus facilitated by *Mmi1* and plays as a decoy to lure *Mmi1*. *Open Biol.* 4, 140022
50 Shimada, T. *et al.* (2003) The fission yeast meiotic regulator Mei2p forms a dot structure in the horse-tail nucleus in association with the *sme2* locus on chromosome II. *Mol. Biol. Cell* 14, 2461–2469
51 Chaumeil, J. *et al.* (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* 20, 2223–2237
52 Plath, K. *et al.* (2002) Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* 36, 233–278
53 Brown, C.J. *et al.* (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349, 38–44
54 Brockdorff, N. *et al.* (1991) Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351, 329–331
55 Brockdorff, N. *et al.* (1992) 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, 515–526
56 Clemson, C.M. *et al.* (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* 132, 259–275
57 Penny, G.D. *et al.* (1996) Requirement for Xist in X chromosome inactivation. *Nature* 379, 131–137
58 Rasmussen, T.P. *et al.* (2001) Expression of Xist RNA is sufficient to initiate macrochromatin body formation. *Chromosoma* 110, 411–420
59 Wutz, A. and Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* 5, 695–705
60 Wutz, A. *et al.* (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* 30, 167–174
61 Kulinski, T.M. *et al.* (2013) Imprinted silencing is extended over broad chromosomal domains in mouse extra-embryonic lineages. *Curr. Opin. Cell Biol.* 25, 297–304
62 Yamasaki, Y. *et al.* (2005) Neuron-specific relaxation of *Igf2r* imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript Air. *Hum. Mol. Genet.* 14, 2511–2520
63 Nagano, T. *et al.* (2008) The *Air* noncoding RNA epigenetically silences transcription by targeting *G9a* to chromatin. *Science* 322, 1717–1720
64 Rinn, J.L. *et al.* (2007) Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* 129, 1311–1323
65 Chu, C. *et al.* (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA–chromatin interactions. *Mol. Cell* 44, 667–678
66 Sun, L. *et al.* (2013) Long noncoding RNAs regulate adipogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3387–3392
67 Carpenter, S. *et al.* (2013) A long noncoding RNA mediates both activation and repression of immune response genes. *Science* 341, 789–792
68 Wang, P. *et al.* (2014) The STAT3-binding long noncoding RNA *lnc-DC* controls human dendritic cell differentiation. *Science* 344, 310–313
69 Orom, U.A. *et al.* (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143, 46–58
70 Margueron, R. and Reinberg, D. (2011) The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343–349
71 Wang, J. *et al.* (2001) Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.* 28, 371–375
72 Mak, W. *et al.* (2002) Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells. *Curr. Biol.* 12, 1016–1020
73 Silva, J. *et al.* (2003) Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of *Eed-Enx1* polycomb group complexes. *Dev. Cell* 4, 481–495
74 da Rocha, S.T. *et al.* (2014) Jarid2 is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome. *Mol. Cell* 53, 301–316
75 Zhao, J. *et al.* (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750–756
76 Cifuentes-Rojas, C. *et al.* (2014) Regulatory interactions between RNA and Polycomb Repressive Complex 2. *Mol. Cell* 55, 171–185
77 Cerase, A. *et al.* (2014) Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2235–2240
78 Brockdorff, N. (2013) Noncoding RNA and Polycomb recruitment. *RNA* 19, 429–442
79 Plath, K. *et al.* (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300, 131–135
80 Schoeftner, S. *et al.* (2006) Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* 25, 3110–3122
81 Kalantry, S. and Magnuson, T. (2006) The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. *PLoS Genet.* 2, e66
82 Zhao, J. *et al.* (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* 40, 939–953
83 Kaneko, S. *et al.* (2014) Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol. Cell* 53, 290–300
84 Yang, Y.W. *et al.* (2014) Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife* 3, e02046
85 Yap, K.L. *et al.* (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
86 Yang, L. *et al.* (2011) ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* 147, 773–788
87 Tsai, M.C. *et al.* (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693
88 Kaneko, S. *et al.* (2013) PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1258–1264
89 Davidovich, C. *et al.* (2013) Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* 20, 1250–1257
90 McHugh, C.A. *et al.* (2014) Methods for comprehensive experimental identification of RNA–protein interactions. *Genome Biol.* 15, 203
91 Bilodeau, S. *et al.* (2009) *SetDB1* contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* 23, 2484–2489
92 Tahiliani, M. *et al.* (2007) The histone H3K4 demethylase SMCX links *REST* target genes to X-linked mental retardation. *Nature* 447, 601–605
93 Cloos, P.A. *et al.* (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev.* 22, 1115–1140
94 Beletskii, A. *et al.* (2001) PNA interference mapping demonstrates functional domains in the noncoding RNA Xist. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9215–9220
95 Sarma, K. *et al.* (2010) Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22196–22201
96 Harrington, L.A. and Greider, C.W. (1991) Telomerase primer specificity and chromosome healing. *Nature* 353, 451–454
97 Schmitz, K.M. *et al.* (2010) Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* 24, 2264–2269
98 Soruco, M.M. *et al.* (2013) The CLAMP protein links the MSL complex to the X chromosome during *Drosophila* dosage compensation. *Genes Dev.* 27, 1551–1556
99 Simon, M.D. *et al.* (2011) The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20497–20502
100 Wang, C.I. *et al.* (2013) Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in *Drosophila*. *Nat. Struct. Mol. Biol.* 20, 202–209
101 Hasegawa, Y. *et al.* (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* 19, 469–476
102 McElroy, K.A. *et al.* (2014) Are we there yet? Initial targeting of the Male-Specific Lethal and Polycomb group chromatin complexes in *Drosophila*. *Open Biol.* 4, 140006
103 Simon, M.D. *et al.* (2013) High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* 504, 465–469

104 Quinn, J.J. *et al.* (2014) Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat. Biotechnol.* 32, 933–940
105 Williams, A. *et al.* (2010) Interchromosomal association and gene regulation in trans. *Trends Genet.* 26, 188–197
106 Denholtz, M. *et al.* (2013) Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* 13, 602–616
107 Noordermeer, D. *et al.* (2014) Temporal dynamics and developmental memory of 3D chromatin architecture at *Hox* gene loci. *Elife* 3, e02557
108 Sunwoo, H. *et al.* (2009) MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* 19, 347–359
109 Chen, L.L. and Carmichael, G.G. (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* 35, 467–478
110 Shevtsov, S.P. and Dundr, M. (2011) Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.* 13, 167–173
111 Splinter, E. *et al.* (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* 25, 1371–1383
112 Lee, J.T. and Jaenisch, R. (1997) Long-range *cis* effects of ectopic X-inactivation centres on a mouse autosome. *Nature* 386, 275–279
113 Lanctot, C. *et al.* (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat. Rev. Genet.* 8, 104–115
114 Spilianakis, C.G. *et al.* (2005) Interchromosomal associations between alternatively expressed loci. *Nature* 435, 637–645
115 Lai, F. *et al.* (2013) Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* 494, 497–501
116 Li, W. *et al.* (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520
117 Orom, U.A. and Shiekhattar, R. (2013) Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* 154, 1190–1193
118 Dundr, M. and Misteli, T. (2010) Biogenesis of nuclear bodies. *Cold Spring Harb. Perspect. Biol.* 2, a000711
119 Misteli, T. (2001) The concept of self-organization in cellular architecture. *J. Cell Biol.* 155, 181–185
120 Kaiser, T.E. *et al.* (2008) *De novo* formation of a subnuclear body. *Science* 322, 1713–1717
121 Misteli, T. (2005) Concepts in nuclear architecture. *Bioessays* 27, 477–487
122 Lewis, J.D. and Tollervey, D. (2000) Like attracts like: getting RNA processing together in the nucleus. *Science* 288, 1385–1389
123 Gonzalez, I. *et al.* (2014) Identification of regulators of the three-dimensional polycomb organization by a microscopy-based genome-wide RNAi Screen. *Mol. Cell* 54, 485–499
124 Schaaf, C.A. *et al.* (2013) Cohesin and polycomb proteins functionally interact to control transcription at silenced and active genes. *PLoS Genet.* 9, e1003560
125 Atchison, M.L. (2014) Function of YY1 in long-distance DNA interactions. *Front. Immunol.* 5, 45
126 Zhang, H. *et al.* (2013) Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell* 13, 30–35
127 Phillips-Cremins, J.E. *et al.* (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295
128 Guttman, M. *et al.* (2010) *Ab initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* 28, 503–510
129 Ponten, F. *et al.* (2009) A global view of protein expression in human cells, tissues, and organs. *Mol. Syst. Biol.* 5, 337
130 Mili, S. and Steitz, J.A. (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 10, 1692–1694
131 Darnell, R.B. (2010) HITS-CLIP: panoramic views of protein–RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA* 1, 266–286
132 Ule, J. *et al.* (2005) CLIP: a method for identifying protein–RNA interaction sites in living cells. *Methods* 37, 376–386
133 Wang, Z. *et al.* (2009) CLIP: construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins *in vivo*. *Methods* 48, 287–293
134 Brimacombe, R. *et al.* (1988) Intra-RNA and RNA–protein cross-linking techniques in *Escherichia coli* ribosomes. *Methods Enzymol.* 164, 287–309
135 Singh, G. *et al.* (2014) RIPiT-Seq: a high-throughput approach for footprinting RNA:protein complexes. *Methods* 65, 320–332