HHS Public Access

Author manuscript

Curr Opin Genet Dev. Author manuscript; available in PMC 2017 May 03.

Published in final edited form as:

Curr Opin Genet Dev. 2016 February ; 36: 73–82. doi:10.1016/j.gde.2016.03.014.

Diverse regulatory interactions of long noncoding RNAs

Tae-Kyung Kim[1] and **Ramin Shiekhattar**[2]

[1]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX, USA

[2]Sylvester Comprehensive Cancer Center, Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA

Abstract

Long noncoding RNAs (lncRNAs) are emerging as important regulators of diverse biological functions. Studies in the past decade indicate that a large number of lncRNAs are enriched in the nucleus and originate from transcriptionally active regulatory elements. These lncRNAs associate with transcription factors and chromatin regulatory elements to fine-tune the transcriptional output of protein coding genes. Importantly, lncRNAs display exquisite tissue specificity in their expression. Understanding how lncRNAs associate with their protein or nucleic acid partners and how they modulate gene expression provides insight into their scope of biological function. This review discusses notable functional properties and mechanisms of action of lncRNAs that have resulted from recent progress made in the field.

LncRNAs cellular localization and maturation

In recent years, long non-coding RNAs (lncRNAs) have been increasingly appreciated as regulatory molecules that play a functional role in diverse cellular processes. Initial observation that lncRNAs predominantly stay in the nucleus after their synthesis and are tightly associated with chromatin suggested their possible role in transcriptional regulation through epigenetic mechanisms [1–4]. Earlier studies used genome-wide co-expression profile analysis between lncRNAs and protein-coding genes to infer repressor or activator functions of lncRNAs, their potential target genes and cellular pathways. Such analysis in four mouse cell types led to the finding that lncRNAs might exert their function in distinct and diverse biological processes such as embryonic stem cell pluripotency, cell proliferation, or neural process [1,2]. Expression of lncRNAs exhibit highly cell-type or developmental stage-specific patterns, and are often dysregulated in a disease state [1,4,5]. LncRNAs are emerging as significant players in diverse aspects of tumorigenesis and metastasis such as DNA damage and cell cycle control (reviewed in [6]). For example, expression analysis of the *HOX* loci identified a systematic variation in expression levels of lncRNAs among normal breast epithelia, primary tumor, and metastases. One of the lncRNAs, *HOTAIR*

Correspondence to: ; Email: Taekyung.kim@utsouthwestern.edu (T.-K. Kim); ; Email: rshiekhattar@med.miami.edu (R. Shiekhattar)

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

showed a strong predictor value of breast cancer metastasis. Overexpression of *HOTAIR* was sufficient to cause increased cancer invasiveness and metastasis [7].

The GENCODE consortium effort on lncRNA profiling has yielded annotation of more than 15,000 lncRNAs present in human cells (http://www.gencodegenes.org/) [4]. Initially identified lncRNAs were deemed to undergo maturation processes such as splicing and polyadenylation, therefore exhibiting the same structural features as protein-coding mRNAs [1]. However, compared to protein-coding mRNAs, they are generally present at much lower levels in the cell with some exceptions and exhibit rather modest level of evolutionary conservation. Adding to its complexity, a rather unique class of lncRNAs, termed enhancer RNAs (eRNAs) has emerged in parallel with conventional lncRNAs described above [8,9]. Enhancer RNAs are divergently transcribed by two independent RNA polymerase II (RNAPII) molecules at individual active enhancers. While the 5′ end of eRNAs are capped, unlike the originally defined lncRNAs, the major population of eRNAs are monoexonic and are not polyadenylated. Perhaps due to these characteristics, transcribed eRNAs are relatively unstable (half-life of ~7 min) and subject to rapid degradation [9,10].

Integrator, a multi-subunit complex associated with RNAPII functions in biogenesis of eRNAs through its catalytic RNA endonuclease activity that mediates 3′-end cleavage of eRNA primary transcripts leading to transcriptional termination [11]. Notably depletion of integrator reduces eRNA induction levels and enhancer-promoter looping, suggesting that proper termination of eRNAs is critical for eRNA biogenesis as well as enhancer function. Although these features of eRNAs would make them distinguishable from originally characterized canonical lncRNAs, recent transcriptome and functional analyses have suggested that many intergenic lncRNAs could also be categorized as eRNAs [12–14]. *De novo* transcriptome assembly of a large set of RNA-seq data covering a wide array of human tissue types identified over 50,000 lncRNAs present at a minimum of at least one copy per cell, which is a significant expansion from the manually catalogued GENCODE lncRNA dataset [12]. These lncRNAs are composed of both poly A+ and poly A− annotated transcripts [4,5,15]. This bimorphic feature of lncRNAs is reproducibly observed across multiple human cell types whereas protein-coding transcripts are strongly enriched in the polyA+ sample [12]. Stringent analysis of a set of intergenic lncRNAs by examining enhancer and promoter-characteristic histone marks (high K4me1/me3 and low K4me1/me3, respectively) for accurately determined transcription initiation sites revealed that intergenic lncRNAs are almost evenly divided between those arising from enhancer-associated (elncRNA) or promoter-associated (plncRNA) elements [14]. For example, a functionally characterized lncRNA, *ncRNA-a3*, displayed similar characteristics to eRNAs as it was shown to originate from a bi-directionally transcribed enhancer of the *TAL1* gene [13,16]. Therefore, some previously annotated lncRNAs appear to originate from functionally active enhancers, resulting in their new classification as eRNAs.

An increasing number of lncRNA and eRNA is shown to have a biological function based on experimental data, which is mainly obtained from loss-of-function or gain-of-function types of analysis (reviewed in [17]). However, we are still far from unambiguous assignment of function for most newly discovered lncRNAs. This effort should also include identification of transcripts previously annotated as lncRNAs that actually encode micropeptides as recent

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genome-wide studies have proposed [18,19]. As the field of lncRNAs matures, additional tools and methodology are being developed to functionally characterize lncRNAs. These include functional validation of lncRNAs *in vivo* by creating KO mice [20–22]. Additional techniques include genome-wide identification of lncRNA interaction loci on chromosomes as well as their binding protein partners using chromatin isolation by RNA purification (ChIRP) [23] and capture hybridization analysis of RNA target (CHART) [24]. One effective way of validating lncRNA functionality is to determine proteins that directly bind to a lncRNA of interest and mechanistically characterize the significance of their interaction, which in principle should provide new molecular insight into the lncRNA role in biology.

Specific and promiscuous binding capacity of lncRNAs

Polycomb repressive complex-2 (PRC2) is a multi-subunit complex that mediates epigenetic silencing during development [25]. One of the core subunits present in the complex, EZH1/2 has methyltransferase activity that critically underlies PRC2 function in gene silencing by generating a repressive histone modification mark, H3K27me3. The subunits of the PRC2 complex have been the most common proteins identified in many studies as RNA-interacting partners associating with thousands of mRNAs and lncRNAs.

A microarray-based RNA immunoprecipitation (RIP) assay for ~1,000 human lncRNAs showed ~ 20% of tested lncRNAs interacting with PRC2 [3]. An early study of X chromosome inactivation (XCI) suggested sequence-specific interactions between lncRNAs and PRC2. PRC2 directly interacts with a 1.6 kb ncRNA *RepA* within *Xist* through its EZH subunit [26]. Furthermore, it was shown that the antisense *Tsix* RNA inhibits this interaction. *RepA* contains short tandem repeats of a 28-nt sequence that folds into two conserved stem-loop structures. The sequence of this repeating unit was sufficient for PRC2 binding. Complementing these studies, a subsequent PRC2 transcriptome study showed the extensive localization of PRC2 near the 5′ end of repressed genes. Many short RNAs associated with paused RNAPII near the 5′ end of the Polycomb target genes (50–200 nt) were proposed to recruit PRC2, resulting in gene repression in *cis* [27]. The interaction was mediated via their stem-loop structures similar to that of *RepA* and SUZ12 subunit of PRC2. Artificial incorporation of a single unit of a two hairpin motif present in the PRC2 bound RNAs in front of the luciferase gene conferred specific binding of luciferase mRNAs with SUZ12 and also mediated a significant repression of luciferase expression. Similarly, another PRC2 transcriptome study in embryonic stem cells (ESCs) using native RIP-Seq determined that nearly 10,000 transcripts are present in various locations (antisense, intergenic, promoter-associate) throughout the genome [28]. An RNA EMSA assay with a few selected RNAs identified from the RIP-Seq additionally demonstrated direct and specific interactions between EZH2 and the stem-loop structure present in the tested RNAs. These studies collectively proposed that a single or tandem short stem-loop structure could serve as a PRC2 binding motif that mediates specific interactions between lncRNAs and PRC2.

However, additional characterizations of the RNA binding specificity both *in vitro* as well as *in vivo* have provided somewhat perplexing results suggesting that a functional PRC2-RNA interaction could occur both in a specific and promiscuous way. The affinity of PRC2 to

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

known specific RNA targets, *HOTAIR* and *RepA* was measured to be in the mid-nanomolar range, but surprisingly a similar affinity was observed for irrelevant RNAs including bacterial mRNA [29]. Such promiscuous PRC2 binding was also suggested *in vivo* by showing that PRC2 associates with both repressed and active genes. This result was in disagreement with an earlier study showing that PRC2 binds *RepA* selectively, with high specificity compared to non-relevant RNA transcripts [30]. A concerted effort toward consensus on the binding specificity and promiscuity of PRC2 for RNA suggested that promiscuous and specific RNA-binding activities of PRC2 *in vitro* are not mutually exclusive and both can be functional [31,32]. Both human and mouse PRC2 complexes bind RNA with mid to low nanomolar affinity although *RepA* showed several fold higher affinity than size-matched irrelevant mRNAs under specific binding conditions. PRC2 binding affinity was proportionally increased with RNA length irrespective of sequence (see also review in [29]).

In an *in vivo* context, other factors influence the intrinsic binding property of PRC2 [32]. Association of JARID2 with PRC2 inhibits its interaction with RNA [30]. Phosphorylation of the EZH2 subunit in a cell cycle dependent manner enhances the RNA binding affinity [33]. ATRX directly promotes the loading of *Xist* to PRC2 and is required for genome-wide PRC2 localization on chromatin [34]. Conversely RNA loading to PRC2 was shown to inhibit the HMT activity of PRC2 [30,35,36]. These allosteric features and the aforementioned promiscuous nature of PRC2-RNA interaction imply that the RNA-guided PRC2 recruitment to chromatin *in vivo* involves rather complex mechanisms.

A prediction that could be inferred from the promiscuous nature of PRC2-RNA interaction is that PRC2 might track along transcriptionally active regions and indeed analysis of EZH2 RIP-seq data sets generated with mouse ESCs revealed broad association of PRC2 in both transcriptionally repressed and active genes [29]. Consistently PRC2 occupies a large fraction of active promoters in mouse ESCs at low levels through direct binding of EZH2 to the 5′ region of nascent RNAs transcribed from the promoter regions [37]. Those promoter regions showing the interaction with PRC2 still exhibit typical active promoter marks (H3K4me3) with little repressive mark (H3K27me3). An emerging model based on these observations is that PRC2 can be targeted to broad regions of chromatin through its promiscuous interaction with various types of RNAs, but its subsequent repressive activity depends on local chromatin environment [32]. PRC2 at active promoters might be released by highly expressed elongating RNAs (decoy) or its HMT activity is suppressed by active histone marks such as H3K4me3 and H3K36me3 [38,39] or by nascent RNA [30,35,36]. Scanning through the genome using this RNA-dependent mechanism allows PRC2 to reach the genes that have been silenced and to deposit repressive H3K27me3 marks, thereby stably maintaining the repressive state. Co-existence of specific and promiscuous interaction capacity of PRC2 and additional allosteric regulation imposed by other factors *in vivo* could explain why many lncRNAs have been shown to interact with PRC2 and support the idea that the promiscuous nature of RNA-protein interaction – without requiring any discernable sequence or structural motifs but with the affinities higher than the non-specific interaction levels [32] – might be a widely utilized molecular mechanism by which many lncRNAs interact with and regulate their cognate target proteins.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Multimodal interaction capacity of lncRNAs

A growing number of lncRNAs show their capacity to interact with more than one protein partner depending on the context (Figure 1). The lncRNA steroid receptor RNA activator (SRA) probably represents a prominent example for having the multimodal capacity to interact with multiple protein partners [40]. Originally identified as a novel modulator of nuclear receptor (NR) function in a ligand-dependent manner [41], SRA interacts directly and indirectly with many other transcription regulators [40]. Notably, SRA forms a complex with the DEAD box RNA helicase p68, which in turn interacts with the MyoD transcription factor to promote muscle gene expression and cellular differentiation [42]. The p68/SRA complex can also contribute to proper insulator function of CTCF by stabilizing the interaction of cohesin with CTCF [43]. In breast cancer cells, SRA can interact with unliganded PR, HP1γ, and LSD1 to form a repressive complex and silence PR target genes [44]. Another intriguing property of SRA is that it has the ability to interact with both the epigenetic repressor and activator, PRC2 and trithorax group (TrxG) complexes, respectively. This binding property is consistent with the finding that some SRA binding sites in human pluripotent stem cells exhibit bivalent domains (H3K27me3-H3K4me3). However association of p68 helicase with SRA tips the balance toward favoring the interaction with TrxG. Therefore, SRA may function as a scaffold to organize multiple factors that regulate gene expression in a context-specific manner.

A lncRNA, *Fendrr* provides another case for dual interaction capacity with PRC2 and TrxG. For proper development of the heart and body wall, *Fendrr* controls expression of several transcriptional regulators in E8.5 embryonic hearts and caudal ends by altering chromatin architecture at the promoter regions of those genes [45]. *Fendrr* could do this by its ability to interact with both PRC2 and WDR5 subunit of TrxG/Mll. The primary role of *Fendrr* is to repress lateral plate mesoderm (LPM) controlling genes by recruiting the PRC2 complex to their promoters. In parallel, *Fendrr* is also involved in the upregulation of a separate set of genes by increasing the level of the activating H3K4me3 mark, thereby modifying the expression level of those genes. However how *Fendrr* can mechanistically accommodate those two complexes with opposing functions has not been understood.

Kcnq1ot1 is a lncRNA transcribed in an antisense orientation with respect to its host gene *Kcnq1* playing an important role in the silencing of eight to ten protein-coding genes spread over a 1 Mb region [46]. It interacts with two different HMTs (G9a and PRC2) as well as DNMT1 to silence both ubiquitously and placental-specific imprinted genes. The 5′ end of *Kcnq1ot1* RNA contains 890 bp silencing domain that harbors several conserved repeats [47]. This domain is specifically required for recruiting DNMT1 to maintain CpG methylation of somatic differentially methylated regions (DMRs) acquired during post-implantation development [48]. This mechanism selectively underlies the maintenance of silencing of ubiquitously imprinted genes [49]. On the other hand, repressive histone modification mediated by G9 and PRC2 was shown to mediate the imprinting of placental-specific imprinted genes [50]. These studies demonstrate how dynamic association of lcnRNA with its various protein partners underlies a lineage-specific transcriptional silencing mechanism for imprinted genes.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Recently, a novel approach called RNA antisense purification followed by quantitative mass spectrometry (RAP–MS) expanded the *Xist* interactome [51]. Three proteins (SHARP, SAF-A and LBR) were required for *Xist*-mediated transcriptional silencing. Identification and additional characterization of these proteins advanced mechanistic understanding of *Xist*-mediated XCI. Briefly, SAF-A was previously shown to interact directly with *Xist* and is required for tethering *Xist* to the inactive X chromosome in differentiated cells [52]. *Xist* interacts directly with SHARP to recruit SMRT to these DNA sites and activates HDAC3 [53]. This cascade of interactions results in gene silencing and chromosome condensation by promoting histone deacetylation as well as RNAPII exclusion. Following the initiation of XCI, *Xist* recruits PRC2 to the X chromosome in an HDAC3-dependent manner to maintain the epigenetic inactive state. Therefore, this study illustrates the usefulness of identifying lncRNA interaction partners in understanding the biological function and the mechanism of action of lncRNAs.

Diverse mechanisms for targeting of lncRNAs to specific genomic loci

An important yet unresolved question is how lncRNAs can find their cognate targets to exert their function. Currently available evidence suggests several possible mechanisms (Figure 2).

1) Recruitment of lncRNA by specific DNA binding proteins

The *roX1* and *roX2* non-coding RNA genes are integral components of the male-specific lethal (MSL) dosage compensation complex in *Drosophila*, which is responsible for increasing transcript levels on the single male X chromosome to equal the transcript levels in XX females [54]. Recruitment of the MSL complex to their entry sites on the X chromosome is mediated by a zinc finger protein, CLAMP (chromatin-linked adaptor for MSL proteins). CLAMP recognizes and binds GA-rich sequences called MSL-recognition elements (MREs), and brings the MSL complex to those sites to initiate the dosage compensation process [23,24,54]. *Dali* is a conserved lncRNA that control a large number of neural differentiation genes *in trans* [55]. It can be recruited to the promoters of target genes primarily by its ability to directly interact with DNMT1, which in turn take *Dali* to distantly located target genes via indirect interactions with several DNA binding proteins.

2) Through the formation of RNA/DNA triple helix

RNA:DNA triple helixes can be formed by Hoogsteen- or reverse Hoogsteen base-pairing between single-stranded RNAs and DNA strands [56]. *De novo* CpG methylation of rRNA genes is mediated by promoter associated RNA (pRNA) complementary to the rDNA promoter. pRNA was shown to form a DNA:RNA triplex which is then specifically recognized by the DNA methyltransferase DNMT3b [57]. During differentiation of lateral mesoderm, a lncRNA, *Fendrr* is targeted to the promoters of *Foxf1* and *Pitx2* genes by forming a triplex structure with double stranded *Foxf1* and *Pitx2* promoter fragments at the complementary region [45]. MEG3 is targeted to the vicinity of the TGF-β pathway genes through the formation of RNA:DNA hybrid triplex structures using its GA-rich sequences. MEG3 then negatively regulates those TGF-β pathway genes by recruiting PRC2 [58]. MEG3-mediated triple helix formation was observed both *in vitro* and *in vivo* and appears to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

occur at a large number of TGF-β pathway genes, suggesting that triplex helix formation might be the major targeting mechanism of MEG3.

R loops are three-stranded nucleic acid structures that are formed by nascent RNA hybridizing to the DNA template, leaving the nontemplate DNA single-stranded [59–61]. In mammalian cells, R loops are abundantly formed at the 5′ ends of genes with G-rich transcripts, as well as near RNAPII pause sites [62–64]. A recent study demonstrates that the R loop can shape the epigenetic landscape and control the differentiation program in ESCs by differentially recruiting two key chromatin-regulatory complexes, Tip60–p400 histone acetyltransferase complex and PRC2 [65]. The genes that form no or low level of R loops are preferential binding sites for PRC2 but are poor Tip60-p400 substrates. Conversely, Tip60–p400 prefers genes that form a high level of R loops, which is not a good substrate for PRC2. This study highlights the importance of the molecular context in which the RNA is presented as a key factor in recruiting a regulatory complex.

3) RNA:RNA interaction

Some lncRNAs form RNA:RNA complexes with other RNA species as a targeting mechanism. The function of *lincRNA-p21* in translation inhibition relies on its ability to interact with target mRNAs, which is promoted by a translational repressor protein, RcK [66]. *EBER2*, an abundant nuclear noncoding RNA expressed by the Epstein-Barr virus (EBV) is specifically targeted to the terminal repeats (TRs) of the latent EBV genome through the base-pairing with nascent transcripts from the TR locus [67]. The *EBER2* then promotes the recruitment of the B cell transcription factor PAX5 to the TR locus. Showing the functional relevance of the *EBER2* targeting mechanism, EBER2 knockdown phenocopies PAX5 depletion and also decreases EBV lytic replication. The *EBER2* guide function of PAX5 to the TRs was also observed in primate herpesvirus CeHV15, suggesting that such a pairing mechanism of a *trans*-acting noncoding RNA might be evolutionarily conserved. Another example is shown by *TINCR*, a lncRNA that controls human epidermal differentiation by a post-transcriptional mechanism [68]. It stabilizes a large number of differentiation mRNAs post-transcriptionally by directly interacting with them through a 25-nucleotide ‘TINCR box’ motif that is strongly enriched in interacting mRNAs. The effect of *TINCR*-mediated stabilization of differentiation mRNAs was also dependent on interaction with the staufen1 (STAU1) protein.

4) LncRNA targeting by enhancer-promoter looping

Chromatin organization in the nucleus can arrange specific lncRNA target genes to be in close proximity to the origin of lncRNA transcription (see also review in [69]). This mechanism would in principle allow extensive interactions between lncRNAs and their target sites located in the same or even different chromosomes. *HOTTIP* is responsible for coordinating activation of *HOXA* genes despite its origin at the distal end of the human *HOXA* cluster [70]. This long-range action is mediated by pre-configured chromosomal looping that brings *HOTTIP* into close proximity of its target genes. *HOTTIP* then brings a member of the Mixed Lineage Leukemia (MLL) family of SET domain-containing lysine methyltransferases, MLL1 by physically interacting with the adaptor protein WDR5 to drive transcription of *HOXA* genes by promoting H3K4me3 modifications. Although several

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

lncRNAs and eRNAs have been shown to play a role in promoting enhancer-promoter looping [17,71–74], the eRNA implicated in the regulation of immediate early genes (IEGs) in neurons appears to act downstream of enhancer-promoter looping. In the study of enhancer function in *Arc* gene induction in neurons, the enhancer-promoter interaction occurs prior to eRNA synthesis in a stimulus-dependent manner and appears to be pre-requisite for eRNA transcription as *Arc* eRNA was only transcribed from WT but not when the *Arc* promoter region is deleted despite the wild-type level of RNAPII binding occurring at the enhancer [8]. Consistently, knockdown of *Arc* eRNA did not affect the enhancer-promoter interaction. Instead the eRNA promotes *Arc* induction by facilitating the release of negative elongation factor NELF from paused RNAPII through its competition with nascent RNAs [10]. Although further study would be necessary to see if this type of eRNA action commonly occurs at many enhancers, the study proposes a model that eRNA function at specific targets can be arranged by chromosomal looping. Transcription of eRNAs generally occurs with a faster kinetics than target protein-coding RNAs, but they are inherently unstable with a observed half-life less than 7.5 min, which would prevent eRNAs from acting distantly from their transcription sites [9,10,75]. Taken together, these findings suggests that the localized abundance of eRNAs in the proximity of target genes might allow effective and specific eRNA action, which is coordinately arranged by multiple mechanisms such as the timing and kinetics of eRNA transcription and inherent RNA stability. A similar mechanism has been proposed for site-specific action of *Tsix* RNA in facilitating locus-specific targeting of CTCF. Newly transcribed *Tsix* RNA selectively recruits CTCF to the site of synthesis but then rapidly turns over to enable its site-specific action [76].

Allosteric regulation by lncRNAs

An increasing number of studies are finding that lncRNAs can allosterically alter the activity of their interacting proteins (Figure 3). An RNA-binding protein, TLS regulates transcription by inhibiting CBP HAT activity in a RNA-dependent manner [77]. The N terminus of TLS possesses a strong inhibitory activity for CBP HAT but binding of TLS C-terminus prevents its inhibitory function. A noncoding RNA ($ncRNA_{CCND1}$) expressed from the 5′ regulatory regions of a TLS target gene, *CCND1* was shown to allosterically modify TLS *in cis* to relieve its auto-inhibitory configuration, thereby repressing *CCND1* expression. lncRNAs can also influence enzymatic activity of chromatin remodelers. *Evf2* is a lncRNA involved in neural development by regulating expression of homeodomain transcription factors DLX5 and DLX6 in the developing mouse forebrain [78]. It forms a complex with transcription DLX homeodomain proteins at the ultraconserved intergenic regions to repress gene expression. Mass spectrometry analysis of the *Evf2*-DLX1 complex revealed the association of the SWI/SNF-related chromatin remodelers Brahma-related gene 1 (BRG1, SMARCA4) and Brahma-associated factor (BAF170, SMARCC2) in the developing mouse forebrain [79]. The association of BRG1 with *Evf2*-DLX1 is mediated by direct interaction with DLX1 but *Evf2* increases BRG1 binding to key *Dlx5/6* enhancers and also inhibits BRG1 ATPase and chromatin remodeling activities, causing gene repression. Additional *in vitro* studies show that both RNA-BRG1 binding and RNA-dependent inhibition of BRG1 ATPase/remodeling activity are rather promiscuous, suggesting that context is a crucial factor in RNA-dependent chromatin remodeling inhibition [79]. Very recently, it was shown

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that lncRNAs transcribed from regulatory elements such as eRNAs and promoter-associated ncRNAs can regulate gene expression *in cis* by stabilizing the interactions between transcription factors (TFs) and the regulatory elements they occupy [80]. Although how mechanistically these ncRNAs can mediate this function is not understood, perturbation of RNA levels affects binding levels of the TF, YY1 across the entire genome, suggesting that stable maintenance of TF binding might be a widespread mechanism contributed by regulatory element-derived lncRNA. It also illustrates a positive feedback loop that reinforces regulatory elements contributing to the stability of gene expression programs.

Future questions for the field

We have come a long way in understanding the diversity of cellular and tissue localization as well as the molecular characteristics of lncRNAs. However, many mechanistic questions remain unanswered. Chief among these is the elucidation of specific molecular determinants of lncRNAs that underlie their precise function. Do lncRNAs contain unique sequence or specific structural bases that govern their association with protein partners or their target DNA/RNA elements? Future experiments using scanning CRISPR/Cas9 mutagenesis of lncRNAs *in vivo* would be an important way to address the structural underpinnings of lncRNAs that result in altered function in specific cellular contexts. It is also clear that additional studies using a variety of organismal models will be necessary in order to understand their conserved evolutionary function as well as their full scope of importance. Using approaches such as genome-wide RNA interference or CRISPR/Cas9 screens in *Caenorhabditis elegans* or Zebrafish for eRNAs and other lncRNAs will be an insightful addition to current studies using mammalian cell culture systems. Furthermore, detailed targeted deletions of lncRNAs and interference with their transcription through insertion of termination sequences will be needed to gain a precise understanding of individual lncRNAs using mouse models. With the advent of new genome-wide technologies the next few years will usher in a renewed understanding of the role of lncRNAs in metazoans.

Acknowledgments

We thank Katie Schaukowitch for her critical reading of the manuscript. We apologize to many investigators whose studies could not be cited due to space limitations. This work was supported by R01NS085418 (T.-K.K.), R01GM078455 and R01GM105754 (R. Shiekhattar) from the National Institute of Health.

References

1••. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009; 458:223–227. This study performs the first genome-wide identification of lncRNAs based on active gene marks, H3K4me3 and H3K36me3. [PubMed: 19182780]

2. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166. [PubMed: 22663078]

3. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009; 106:11667–11672. [PubMed: 19571010]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012; 22:1775–1789. [PubMed: 22955988]

5. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, et al. Landscape of transcription in human cells. Nature. 2013; 488:101–108.

6. Sahu A, Singhal U, Chinnaiyan AM. Long noncoding RNAs in cancer: from function to translation. Trends Cancer. 2015; 1:93–109. [PubMed: 26693181]

7. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010; 464:1071–1076. [PubMed: 20393566]

8••. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465:182–187. This study and [9] describe the first genome-wide evidence for eRNA transcription activity at functionally active enhancers. [PubMed: 20393465]

9. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei C-L, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 2010; 8:e1000384. [PubMed: 20485488]

10•. Schaukowitch K, Joo JY, Liu X, Watts JK, Martinez C, Kim TK. Enhancer RNA facilitates NELF release from immediate early genes. Mol Cell. 2014; 56:29–42. This study demonstrates that eRNAs can specifically function in the early transcription elongation stage by facilitating the release of negative elongation factor (NELF). Authors also propose that local abundance near target genes would be critical for eRNA function and might be determined by tightly controlled eRNA synthesis and its intrinsic instability. [PubMed: 25263592]

11••. Lai F, Gardini A, Zhang A, Shiekhattar R. Integrator mediates the biogenesis of enhancer RNAs. Nature. 2015; 525:399–403. In this study, authors demonstrate that integrator complex, in addition to its previously known role in the 3′ end processing of small nuclear RNAs, is critical for the biogenesis of eRNAs. Integrator is recruited to the enhancers in a stimulus-dependent manner and cleaves the 3′ end of eRNA transcripts leading to termination. Interestingly, functional depletion of Integrator subunits diminishes the signal-dependent induction of eRNAs and abrogates stimulus-induced enhancer-promoter chromatin looping. [PubMed: 26308897]

12. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet. 2013; 9:e1003569. [PubMed: 23818866]

13. Vucicevic D, Corradin O, Ntini E, Scacheri PC, Orom UA. Long ncRNA expression associates with tissue-specific enhancers. Cell Cycle. 2015; 14:253–260. [PubMed: 25607649]

14. Marques AC, Hughes J, Graham B, Kowalczyk MS, Higgs DR, Ponting CP. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. Genome Biol. 2013; 14:R131. [PubMed: 24289259]

15. Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science. 2005; 308:1149–1154. [PubMed: 15790807]

16••. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010; 143:46–58. In this study, authors functionally analyzed a collection of GENCODE annotated lncRNAs and report the observation that over 70% of extragenic lncRNAs function to enhance the expression of their neighboring genes. This is a first report of lncRNAs functioning in a similar capacity as DNA enhancer elements. [PubMed: 20887892]

17. Kim TK, Shiekhattar R. Architectural and Functional Commonalities between Enhancers and Promoters. Cell. 2015; 162:948–959. [PubMed: 26317464]

18. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. Nat Rev Genet. 2014; 15:193–204. [PubMed: 24514441]

19. Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES, Vejnar CE, Lee MT, Rajewsky N, Walther TC, et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. EMBO J. 2014; 33:981–993. [PubMed: 24705786]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

20. Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife. 2013; 2:e01749. [PubMed: 24381249]
21. Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC, Gingeras TR, Haerty W, et al. Considerations when investigating lncRNA function in vivo. Elife. 2014; 3:e03058. [PubMed: 25124674]
22. Kohtz JD. Long non-coding RNAs learn the importance of being in vivo. Front Genet. 2014; 5:45. [PubMed: 24624134]
23•. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell. 2011; 44:667–678. This study and [24] describe similar genome-wide methodologies to determine the target sites of lncRNAs. [PubMed: 21963238]
24. Simon MD, Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, Borowsky ML, Kuroda MI, Kingston RE. The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A. 2011; 108:20497–20502. [PubMed: 22143764]
25. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011; 469:343–349. [PubMed: 21248841]
26. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008; 322:750–756. [PubMed: 18974356]
27. Kanhere A, Viiri K, Araujo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, et al. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. Mol Cell. 2010; 38:675–688. [PubMed: 20542000]
28. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M, Lee JT. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell. 2010; 40:939–953. [PubMed: 21172659]
29•. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. Nat Struct Mol Biol. 2013; 20:1250–1257. This study provides several lines of evidence that suggest promiscuous nature of RNA binding to PRC2. Authors measure the binding constants of human PRC2 to various RNAs by quantitative EMSA and find submicromolar affinity for all tested RNAs including irrelevant transcripts. The study also showed that PRC2 binds not only repressed genes but also some active genes, further supporting the promiscuous nature of RNA binding. [PubMed: 24077223]
30•. Cifuentes-Rojas C, Hernandez AJ, Sarma K, Lee JT. Regulatory interactions between RNA and polycomb repressive complex 2. Mol Cell. 2014; 55:171–185. This study illustrates the dynamic regulatory interaction between PCR2 and RNA. Upon binding, RNA inhibits EZH2's catalytic activity but JARID2 association relieves the inhibition by weakening PRC2's binding to RNA. In contrast to [29] and [37], this study provides evidence that is in favor of PRC2's ability to effectively discriminate between specific and nonspecific RNAs. [PubMed: 24882207]
31••. Davidovich C, Wang X, Cifuentes-Rojas C, Goodrich KJ, Gooding AR, Lee JT, Cech TR. Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. Mol Cell. 2015; 57:552–558. In a concerted effort to obtain consensus on the model of PRC2-RNA interaction, this study demonstrates that promiscuous and specific RNA-binding activities of PRC2 are not mutually exclusive. Authors find that RNA length increases binding affinities irrespective of sequence, and that various longer-length RNAs can bind PRC2. [PubMed: 25601759]
32•. Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. RNA. 2015; 21:2007–2022. This excellent review discusses in details about the models for the recruitment of PRC2 to chromatin by RNAs as well as systematic and practical considerations for studying the nature of protein-RNA interaction. [PubMed: 26574518]
33. Kaneko S, Li G, Son J, Xu CF, Margueron R, Neubert TA, Reinberg D. Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev. 2010; 24:2615–2620. [PubMed: 21123648]
34. Sarma K, Cifuentes-Rojas C, Ergun A, Del Rosario A, Jeon Y, White F, Sadreyev R, Lee JT. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. Cell. 2014; 159:869–883. [PubMed: 25417162]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

35. Herzog VA, Lempradl A, Trupke J, Okulski H, Altmutter C, Ruge F, Boidol B, Kubicek S, Schmauss G, Aumayr K, et al. A strand-specific switch in noncoding transcription switches the function of a Polycomb/Trithorax response element. Nat Genet. 2014; 46:973–981. [PubMed: 25108384]

36. Kaneko S, Son J, Bonasio R, Shen SS, Reinberg D. Nascent RNA interaction keeps PRC2 activity poised and in check. Genes Dev. 2014; 28:1983–1988. [PubMed: 25170018]

37•. Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat Struct Mol Biol. 2013; 20:1258–1264. This study, together with [33, 36] collectively demonstrates promiscuous binding activity of PRC2 that leads to its occupancy in both repressed and active genes. PAR-CLIP analysis reveals that EZH2 directly binds the 5′ region of nascent RNAs, but its HMT activity is inhibited by RNA, resulting in a decreased level of H3K27me3 at those transcriptionally active genes. Based on these findings, authors propose a model that PRC2 activity is poised and in check. [PubMed: 24141703]

38. Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J Biol Chem. 2011; 286:7983–7989. [PubMed: 21239496]

39. Schmitges FW, Prusty AB, Faty M, Stutzer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol Cell. 2011; 42:330–341. [PubMed: 21549310]

40. Colley SM, Leedman PJ. Steroid Receptor RNA Activator - A nuclear receptor coregulator with multiple partners: Insights and challenges. Biochimie. 2011; 93:1966–1972. [PubMed: 21807064]

41. Lanz RB, McKenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, Tsai MJ, O'Malley BW. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell. 1999; 97:17–27. [PubMed: 10199399]

42. Caretti G, Schiltz RL, Dilworth FJ, Di Padova M, Zhao P, Ogryzko V, Fuller-Pace FV, Hoffman EP, Tapscott SJ, Sartorelli V. The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. Dev Cell. 2006; 11:547–560. [PubMed: 17011493]

43. Yao H, Brick K, Evrard Y, Xiao T, Camerini-Otero RD, Felsenfeld G. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. Genes Dev. 2010; 24:2543–2555. [PubMed: 20966046]

44. Vicent GP, Nacht AS, Zaurin R, Font-Mateu J, Soronellas D, Le Dily F, Reyes D, Beato M. Unliganded progesterone receptor-mediated targeting of an RNA-containing repressive complex silences a subset of hormone-inducible genes. Genes Dev. 2013; 27:1179–1197. [PubMed: 23699411]

45. Grote P, Wittler L, Hendrix D, Koch F, Wahrisch S, Beisaw A, Macura K, Blass G, Kellis M, Werber M, et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. Dev Cell. 2013; 24:206–214. [PubMed: 23369715]

46. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell. 2008; 32:232–246. [PubMed: 18951091]

47. Mohammad F, Pandey RR, Nagano T, Chakalova L, Mondal T, Fraser P, Kanduri C. Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. Mol Cell Biol. 2008; 28:3713–3728. [PubMed: 18299392]

48. Mohammad F, Mondal T, Guseva N, Pandey GK, Kanduri C. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development. 2010; 137:2493–2499. [PubMed: 20573698]

49. Mohammad F, Pandey GK, Mondal T, Enroth S, Redrup L, Gyllensten U, Kanduri C. Long noncoding RNA-mediated maintenance of DNA methylation and transcriptional gene silencing. Development. 2012; 139:2792–2803. [PubMed: 22721776]

50. Lewis A, Reik W. How imprinting centres work. Cytogenet Genome Res. 2006; 113:81–89. [PubMed: 16575166]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

51•. McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 2015; 521:232–236. In this study, authors use a quantitative mass spectrometry approach to identify novel interacting partners of *Xist*. Characterization of these associating proteins provides a model for how *Xist* can orchestrate transcriptional silencing on the X chromosome. [PubMed: 25915022]

52. Hasegawa Y, Brockdorff N, Kawano S, Tsutui K, Tsutui K, Nakagawa S. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell. 2010; 19:469–476. [PubMed: 20833368]

53. Shi Y, Downes M, Xie W, Kao HY, Ordentlich P, Tsai CC, Hon M, Evans RM. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev. 2001; 15:1140–1151. [PubMed: 11331609]

54. Soruco MM, Chery J, Bishop EP, Siggers T, Tolstorukov MY, Leydon AR, Sugden AU, Goebel K, Feng J, Xia P, et al. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. Genes Dev. 2013; 27:1551–1556. [PubMed: 23873939]

55. Chalei V, Sansom SN, Kong L, Lee S, Montiel JF, Vance KW, Ponting CP. The long non-coding RNA Dali is an epigenetic regulator of neural differentiation. Elife. 2014; 3:e04530. [PubMed: 25415054]

56. Morgan AR, Wells RD. Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. J Mol Biol. 1968; 37:63–80. [PubMed: 5760495]

57. Schmitz KM, Mayer C, Postepska A, Grummt I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev. 2010; 24:2264–2269. [PubMed: 20952535]

58. Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, Mitra S, Mohammed A, James AR, Hoberg E, et al. MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. Nat Commun. 2015; 6:7743. [PubMed: 26205790]

59. Drolet M, Phoenix P, Menzel R, Masse E, Liu LF, Crouch RJ. Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. Proc Natl Acad Sci U S A. 1995; 92:3526–3530. [PubMed: 7536935]

60. Thomas M, White RL, Davis RW. Hybridization of RNA to double-stranded DNA: formation of R-loops. Proc Natl Acad Sci U S A. 1976; 73:2294–2298. [PubMed: 781674]

61. Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 2014; 28:1384–1396. [PubMed: 24990962]

62. Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell. 2012; 45:814–825. [PubMed: 22387027]

63. Ginno PA, Lim YW, Lott PL, Korf I, Chedin F. GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. 2013; 23:1590–1600. [PubMed: 23868195]

64. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell. 2011; 42:794–805. [PubMed: 21700224]

65. Chen PB, Chen HV, Acharya D, Rando OJ, Fazzio TG. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nat Struct Mol Biol. 2015; 22:999–1007. [PubMed: 26551076]

66. Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JL, De S, Huarte M, Zhan M, Becker KG, Gorospe M. LincRNA-p21 suppresses target mRNA translation. Mol Cell. 2012; 47:648–655. [PubMed: 22841487]

67. Lee N, Moss WN, Yario TA, Steitz JA. EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. Cell. 2015; 160:607–618. [PubMed: 25662012]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

68. Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature. 2013; 493:231–235. [PubMed: 23201690]

69. Quinodoz S, Guttman M. Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. Trends Cell Biol. 2014; 24:651–663. [PubMed: 25441720]

70••. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 2011; 472:120–124. In this study, authors demonstrate the function of lncRNA *HOTTIP* in activation of several *HOXA* genes by a long range interaction in vivo. *HOTTIP* exerts its function by directly interacting with the adaptor protein WDR5 directly and recruiting WDR5/MLL complexes across *HOXA*, driving H3K4me3 and gene transcription. This study also illustrates spatial proximity induced by chromosomal looping underlies the specificity of HOTTIP function. [PubMed: 21423168]

71••. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature. 2013; 494:497–501. This study and [16] demonstrate enhancer-like function of canonical lncRNAs. Mechanistic study further reveals that these activating lncRNAs positively regulate gene activation by binding to Mediator and thereby promoting the enhancer-promoter interaction. Together with studies described in [72,73], these studies establish that facilitation of long-range chromosomal interactions is a mechanism by which lncRNA can exert its function in gene expression. [PubMed: 23417068]

72. Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, et al. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature. 2013; 498:511–515. [PubMed: 23728303]

73. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013; 498:516–520. [PubMed: 23728302]

74. Hsieh CL, Fei T, Chen Y, Li T, Gao Y, Wang X, Sun T, Sweeney CJ, Lee GS, Chen S, et al. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc Natl Acad Sci U S A. 2014; 111:7319–7324. [PubMed: 24778216]

75•. Arner E, Daub CO, Vitting-Seerup K, Andersson R, Lilje B, Drablos F, Lennartsson A, Ronnerblad M, Hrydziuszko O, Vitezic M, et al. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science. 2015 This study reports an intriguing observation made from a large number of transcriptome datasets that enhancer transcription is the earliest transcriptional response during cellular differentiation or activation.

76. Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, Colognori D, Jeon Y, Szanto A, del Rosario BC, Pinter SF, et al. Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. Mol Cell. 2015; 57:361–375. [PubMed: 25578877]

77. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK, Kurokawa R. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature. 2008; 454:126–130. [PubMed: 18509338]

78. Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev. 2006; 20:1470–1484. [PubMed: 16705037]

79. Cajigas I, Leib DE, Cochrane J, Luo H, Swyter KR, Chen S, Clark BS, Thompson J, Yates JR 3rd, Kingston RE, et al. Evf2 lncRNA/BRG1/DLX1 interactions reveal RNA-dependent inhibition of chromatin remodeling. Development. 2015; 142:2641–2652. [PubMed: 26138476]

80••. Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, Jangi M, Giallourakis CC, Sharp PA, Young RA. Transcription factor trapping by RNA in gene regulatory elements. Science. 2015; 350:978–981. This study demonstrates that the RNAs transcribed from the cis-regulatory regions such as eRNAs and promoter-associated ncRNAs could have a modest but important contribution to the maintenance of certain TFs at gene regulatory elements. [PubMed: 26516199]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

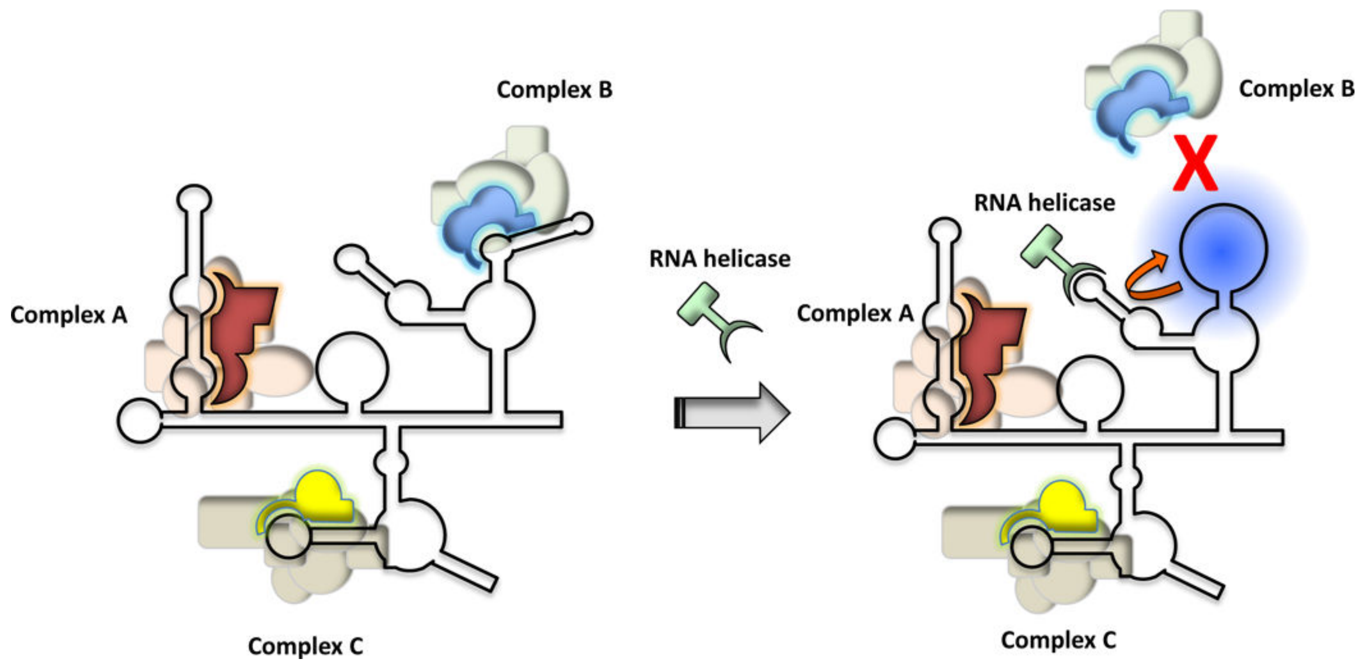

Figure 1. Multimodal interaction capacity of lncRNAs

Several lncRNAs including *HOTAIR*, *SRA*, and *Fendrr* can function as a scaffold organizing multiple proteins or protein complexes (shown in the diagram as complexes A, B, and C). Such multimodal interactions may not be static but rather dynamically regulated depending on the genomic context in which lncRNAs play a role. The complexes shown here can be PRC2 and TrxG for *Fendrr*; PRC2, TrxG, CTCF, and PR for *SRA*; PRC2, DNMT1, and G9a for *Kcnq1ot1*; SHARP, SAF-A and LBR for *Xist* as described in the text.

While *SRA* can interact with both PRC2 and TrxG, association of p68 helicase causes preferential binding of SRA to TrxG, which might be induced by p68-mediated alteration of SRA secondary structure. Note that although the diagram shows that RNA helicase action triggers release of complex B, it might act positively in other context promoting the association of protein complexes with RNA.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

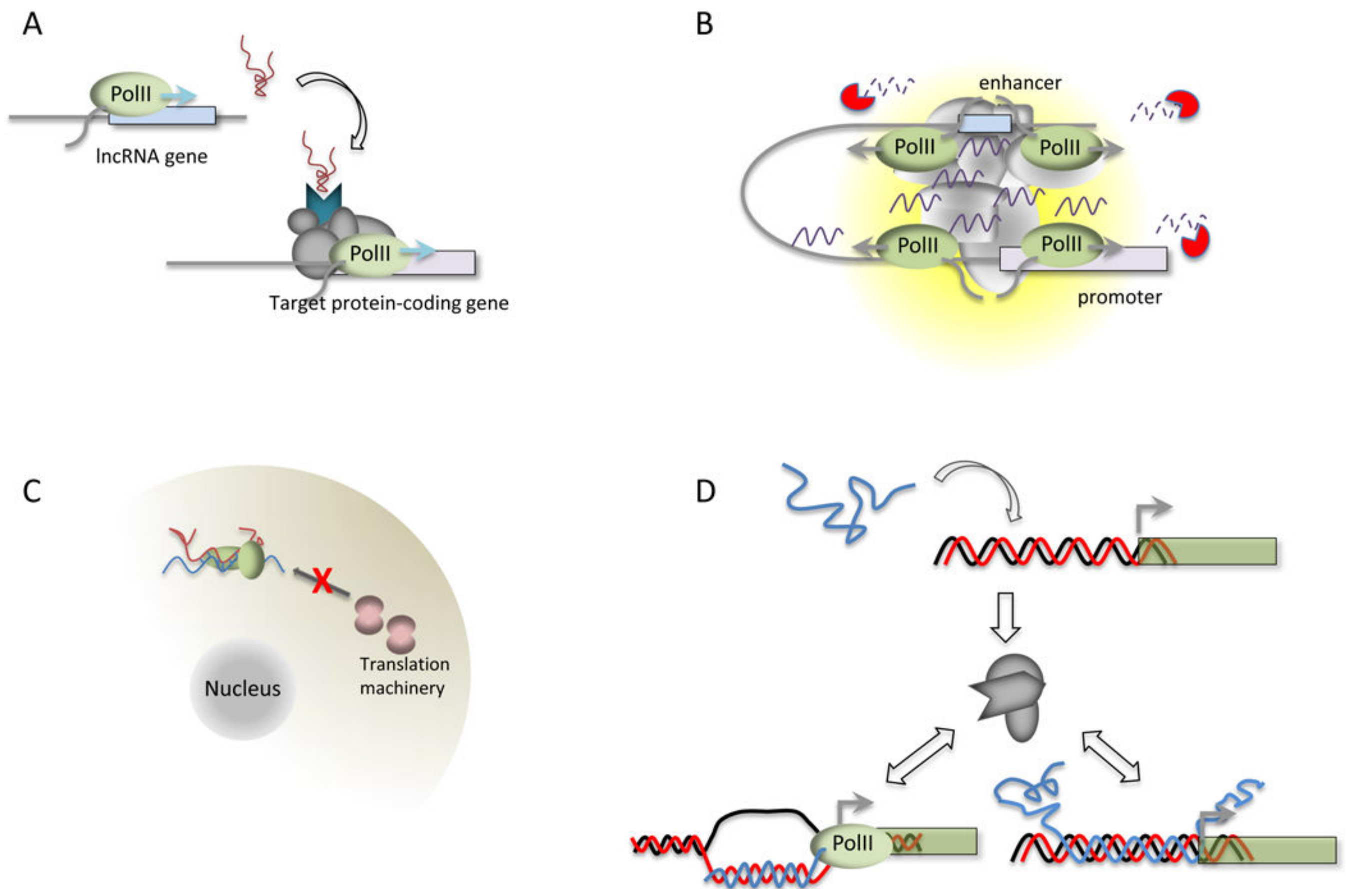

Figure 2. Various targeting mechanism of lncRNAs

A) Recruitment of lncRNA by specific DNA binding proteins. As part of the male-specific lethal (MSL) dosage compensation complex in *Drosophila, roX2* can be recruited to its target genes by CLAMP (chromatin-linked adaptor for MSL proteins). *Dali* can be brought to the promoter region of its target by interacting with DNMT1.

B) Targeting by chromosomal arrangement. *HOTTIP* and *Arc* eRNAs are placed in proximity to their target genes by a pre-configured enhancer-promoter looping mechanism. Targeted degradation of lncRNAs before diffusing away from their origin might be a mechanism to prevent lncRNAs from acting on nonspecific target genes.

C) RNA:RNA interaction. *lincRNA-p21* can function as a translation inhibitor by forming a duplex with target mRNA in the cytoplasm. RNA-RNA base-pairing also occurs in the nucleus as shown by *EBER2*.

D) Formation of RNA/DNA triple helix. Formation of RNA:DNA triplex or R loop between DNA region of target genes (e.g., promoter) and lncRNAs is not only a targeting mechanism for lncRNA (e.g., *MEG3*) but also an allosteric regulation mechanism (e.g., Tip60–p400)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

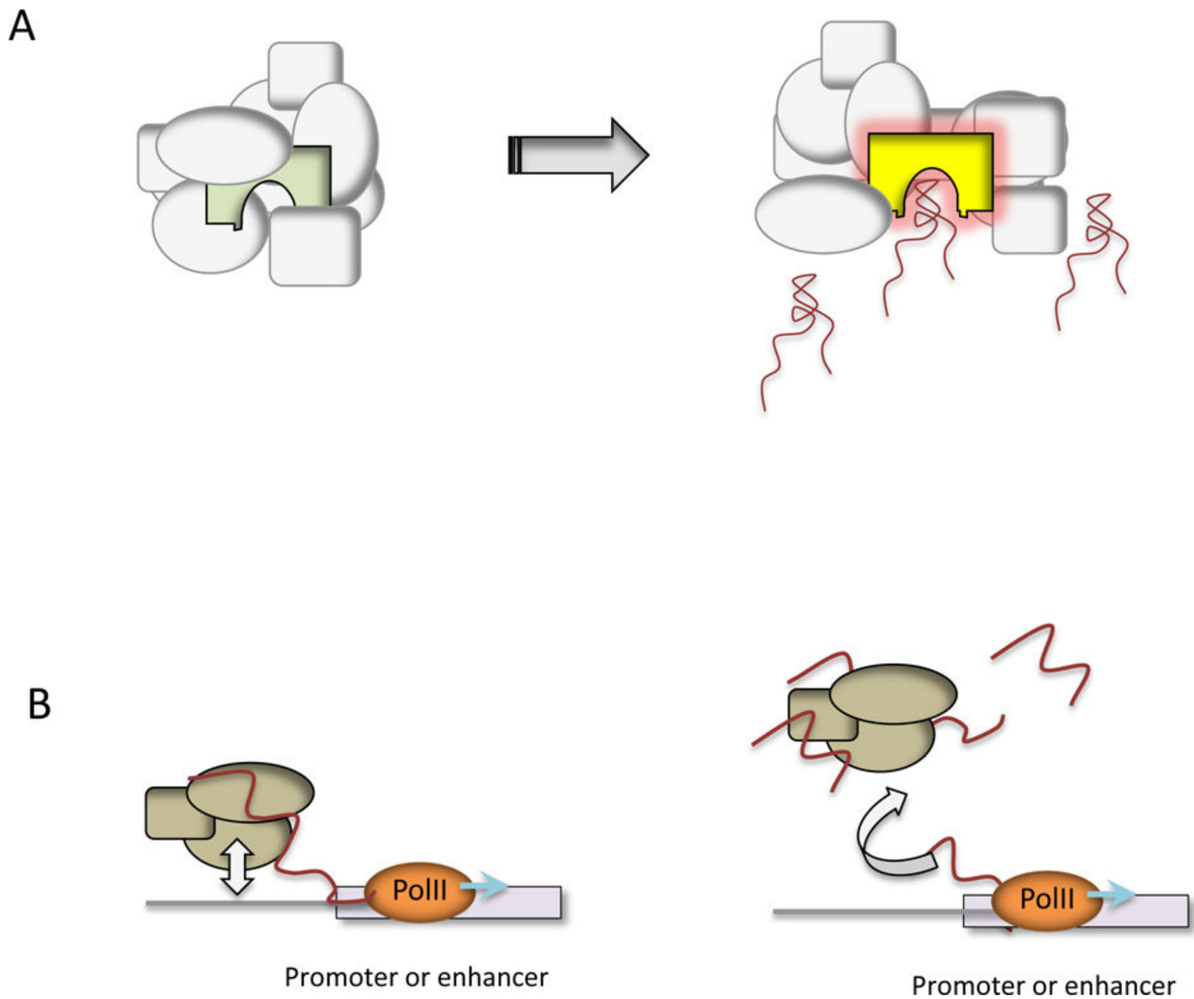

Figure 3. Allosteric regulation by lncRNAs

A) Enzyme activity of epigenetic regulators such as CBP acetyltransferase and the SWI/SNF-related chromatin remodeling complex can be allosterically regulated by lncRNA action. A lncRNA, *Evf2* binds and inhibits BRG1 ATPase activity of the SWI/SNF complex, causing repression of its target genes. $ncRNA_{CCND1}$ indirectly inhibits CBP HAT activity by allosterically activating CBP inhibitor protein, TLS *in cis*.

B) Nascent transcripts attached to RNAPII near the promoter or enhancer can contribute to the stable maintenance of certain transcription factors (e.g., YY1) at their cognate binding sites in *cis*. Alternatively liberated RNAs can facilitate TF release as shown by Arc eRNA.