Published in final edited form as:

FEBS J. 2015 May ; 282(9): 1647–1657. doi:10.1111/febs.13101.

The many faces of long noncoding RNAs

Alessandro Gardini and Ramin Shiekhataar

Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Graphical abstract

Over the past few years the field of noncoding RNAs has grown from a niche for geneticists into a prominent domain of mainstream biology. Advances in genomic technologies have provided a more comprehensive view of the mammalian genome, improving our knowledge of regions of the genome devoid of protein coding potential. A large body of evidence supports the proposal that noncoding RNAs account for a large portion of the transcriptional output of any given cell and tissue type. This review will delve into the biogenesis and function of long noncoding RNAs. We will discuss our current understanding of these molecules as major chromatin players and explore future directions in the field.

The rise of long noncoding RNAs

Long noncoding RNAs (lncRNAs) have been traditionally associated with a handful of transcripts such as *Kcnq1ot1* or *Xist*, implicated in the biological processes involving dosage compensation. However, recent results spurred from technological advances in DNA

Correspondence to: rshiekhataar@med.miami.edu (R. Shiekhataar).

Author Information

The authors declare no competing financial interests.

sequencing indicated that a major portion of the mammalian genome is being transcribed and protein-coding sequences only account for a minority of a cell's transcriptional output [1-3]. These early studies began to generate new interest into what was referred to as 'junk DNA' and were followed by a more comprehensive analysis of the genome upon completion of the ambitious ENCODE project in 2012 [4-6]. The ENCODE consortium generated a large amount of data corresponding to the genome-wide residence of DNA binding factors, chromatin and DNA modifications, which resulted in the compilation of atlas of regulatory elements for both the human and mouse genomes. Among these elements were thousands of sites producing transcripts, which encoded RNA without protein coding potential. Interestingly, the genomic architecture of a noncoding RNA locus was described to be similar to that of a protein-coding gene, consisting of a defined transcriptional start site and an exon-intron structure. The proximal promoter region of such noncoding transcripts boasted all the classic features of coding genes such as a narrow and distinct peak of RNAPII, a nucleosome free region and enrichment in H3K4me3. There were several thousand of noncoding loci across the mammalian genome, about 9,000 of which were reported in human cells based on manually curated annotations [7]. Initial studies in mice focused on about 1,000 predicted lncRNAs from differentiated cells and ES cells [8]. This number has almost tripled according to the most recent estimates in ES cells, and could soar to as many as 20,000 noncoding transcripts in the developing brain [9].

RNA matters

To assess the functional role of lncRNAs, it is critical to rule out any possible protein coding potential. The coding potential of lncRNAs, as measured by GENEID or by codon substitution frequency, was estimated to be far lower than in protein coding genes and quite similar to that of ancestral repeats [8, 10]. Indeed some candidate lncRNAs were examined in an in-vitro translation assay where no polypeptides were detected [10]. Surprisingly, a genome wide ribosome profiling in mouse cells (Ribo-seq) revealed that a large portion of the annotated lncRNAs were exported to the cytoplasm and engaged in translation by elongating ribosomes [11]. The observation that a given lncRNA is associated with ribosomes does not provide direct evidence of the synthesis of a biologically relevant polypeptide. Indeed, later it was determined that ribosome profiling displayed by lncRNAs does not differ from the ribosome association pattern exhibited by well-established ncRNA molecules, such as small nucleolar RNAs or telomerase RNA [12]. While most experimental results point to the lncRNAs lack of protein coding potential, it is possible that a small set of lncRNAs encode short peptides that escape traditional proteomics techniques. There are examples of RNAs centrally implicated in *Drosophila* development and morphogenesis that were erroneously classified as noncoding [13, 14]. These RNAs were later found to encode for short peptides (10-30 amino acids).

Could such a circumstance apply to a much larger group of lncRNAs? The answer was provided following a large unbiased proteomics screening, which also included the analysis of short peptides. Peptidomics is a relatively new and rapidly evolving field that aims at characterizing the entire repertoire of peptides in a biological sample. Using this approach, it was determined that 53 peptides overlapped with known lncRNAs system [15]. Although substantial, this number only represented 0.4% of all lncRNAs detected by RNA-seq in the

same biological system [15]. Similarly, a large proteomic screening in human tissues identified peptides from 400 lncRNAs out of more than 30,000 previously annotated transcripts [16]. While it might be expected that deeper sequencing of the proteomes in the next few years may lead to a large number of lncRNAs with short reading frames, there is compelling evidence that the majority of lncRNAs do not encode for short peptides [15, 16].

Familiar features at lncRNA loci

The genomic structure comprising the promoter region and the transcription start sites of lncRNA are similar to that of protein coding genes. This also applies to transcripts that map to distal regulator elements (enhancers), since enhancers and their corresponding promoters (proximal elements) are quite similar. lncRNA promoters display a significant degree of conservation across mammals and the proximal promoter is a nucleation site for the RNAPII machinery [7, 8]. In fact, ChIP and ChIP-seq experiments show enrichment of the core RNA polymerase II and the general transcription factors at the 5' end of noncoding genes [7, 17]. The chromatin architecture of these loci are defined by a peak of H3K4me3 surrounding the transcription start site and a larger domain enriched in H3K36me3, which encompasses the entire gene body. These signatures were observed both in human and mouse cells [7, 8], and were initially used to define a set of lncRNAs in mouse ES cells [8].

Initial reports indicated that lncRNAs are polyadenylated and spliced [7, 8, 10, 18], although, less efficiently than protein coding genes [19]. While this may be the case for several lncRNAs, there is a large number of noncoding transcripts that are monoexonic, non-polyadenylated and originate from distal regulatory regions [20]. Moreover, for those lncRNAs that are spliced, the exon-intron structure of lncRNAs is considerably simpler than that displayed at protein coding loci. Finally, lncRNAs are capped by 5' methyl-guanosine and show a slightly smaller cap analysis gene expression (CAGE) tag coverage than protein coding genes [7].

The spliced and polyadenylated lncRNAs display two major differences with messenger RNAs: i) their exon-intron structure is relatively simpler with nearly half of lncRNAs only bearing two exons [7] and ii) while they display exquisite pattern of tissue specificity, their expression levels are significantly lower than protein coding genes [7, 21]. Median expression levels of lncRNAs (measuring steady-states of transcripts) are roughly 10 times lower than those of mRNAs (with a few notable exceptions such as Malat1, Neat1 and Xist that are highly abundant) [7]. Importantly, lncRNAs show a prominent tissue specificity [21], as previously mentioned, suggestive of a critical role for these molecules in developmental control.

Are lncRNAs evolutionarily conserved?

While initial studies focused on mouse and human noncoding repertoires, additional reports identified lncRNAs in *Xenopus*, Zebrafish, *Drosophila* and *C. Elegans* [22-25]. Although lncRNAs were not subjected to the same evolutionary pressures as protein coding genes [8, 10], their exons showed a significant degree of conservation. Detailed analysis has revealed that lncRNA exons are more conserved than intergenic regions devoid of transcripts. Interestingly, lncRNA promoters show a greater evolutionary conservation compared to

their downstream noncoding transcripts [7, 8, 10]. It must be noted that using conventional approaches it has been difficult to discern homologous regions in noncoding RNAs [18]. Importantly, in many cases while the specific sequences of lncRNA may have diverged there is often evidence of the presence of lncRNAs in similar syntenic regions across multiple organisms [18]. This finding is consistent with the stronger extent of conservation displayed at lncRNA promoters, and conveys the idea that lncRNA loci are transcribed regulatory elements governed by sequence specific transcription factors, where the RNA sequence is allowed far more flexibility than that of protein coding genes.

In order to gain greater insight into the function of lncRNAs, it will be important to understand the underlying structural features that allow lncRNAs to mediate their biological effects. lncRNAs are known to form secondary structures that allow for their proper 3'-end processing [26]. Much like proteins, where structural features are quite conserved across evolution, it is likely that while lncRNA primary nucleotide sequences may have diverged, their structural elements have remained constant in higher eukaryotes. Taken together, our current view on the secondary and tertiary structure of these transcripts remains quite limited and will require the development of additional methodologies to allow for a more detailed elucidation of lncRNA structure.

Repressive functions of lncRNAs

The best functional characterizations of lncRNAs are in the epigenetic phenomena of X inactivation and imprinting, in which the lncRNA triggers the transcriptional silencing of the neighboring genes [27, 28]. *HOTAIR* is an example of a lncRNA with repressive functions [29]. While *HOTAIR* is embedded within the *HOXC* cluster on chromosome 12 and encodes for a transcript of about 2 kilobases long, its depletion led to activation of multiple *HOXD* genes on chromosome 2. Such a de-repression of *HOXD* genes was accompanied by loss of H3K27me3 and Polycomb Repressive Complex (PRC2) binding across a major portion of the *HOXD* cluster. It was postulated that *HOTAIR* functions to repress genes in a *trans* fashion by facilitating recruitment of Suz12 to its target *HOX* genes [29]. Further structure/function analysis mapped *HOTAIR* association with PRC2 to the 5' end of the transcript, while the 3' end exhibited binding to LSD1 [30]. Therefore, *HOTAIR* was proposed to act as a modular scaffold coordinately regulating PRC2 methyltransferase and LSD1 demethylase activities. This latter report extended *HOTAIR* scope of function beyond the previously described *HOXD* cluster, increasing the number of target loci to a broad set of genes that are co-occupied by LSD1 and PRC2 [30].

Additional lncRNAs were reported to act as repressors with a similar mechanism as *HOTAIR*, culminating in the recruitment of PRC2. For instance, *Braveheart*, a mouse specific lncRNA, is required for cardiac differentiation [31] and regulates gene expression programs at several stages of the cardiovascular development via direct binding to Suz12. Additionally, *Xist* and other noncoding RNAs were shown to bind JARID2, a component of PRC2 complex [32, 33]. Notably, *Xist* may also tether PRC2 to chromatin through interaction with YY1, establishing a role for lncRNAs in modulating the effect of a sequence specific transcription factor [34]. In other cases the repressive mechanism entails the disruption of chromatin insulation mediated by CTCF and the Cohesin complex [35], or

involves association of the lncRNA (such as the case of lncRNA-p21) with the repressor hnRNP-K to counteract p53-dependent activation [36]. In summary, there are now several reports of lncRNAs acting as transcriptional co-repressors and multiple mechanistic scenarios have been described. The prevailing hypothesis depicts an association of lncRNA with different subunits of the PRC2 complex, with variable target selectivity. Affinity of PRC2 for different lncRNAs may be partly explained by a promiscuous affinity of EZH2 for long RNAs (>300bp) [37, 38], including nascent transcripts [39]. Broad affinity of PRC2 for RNA is a two-edged sword, since it may help recruit Polycomb proteins to chromatin, but would require additional events to overcome the RNA-dependent inhibition of PRC2 catalytic activity [38, 39].

The overarching hypothesis for repressive noncoding RNAs contends that a number of multiprotein complexes involved in transcriptional repression may coordinate their function by associating with lncRNAs. Indeed, the characterization of physical interactions between several lncRNAs and chromatin modifying complexes has generated a great deal of excitement. However, the current biochemical assays suffer from elevated backgrounds and large false positive rates and one cannot rule out the possibility that some lncRNA-protein interactions may be indirect. Techniques of RNA-protein Cross-Linking followed by Immunoprecipitation (CLIP) have dramatically improved over the past few years, delivering higher resolution mapping of RNA-protein interactions as in the case of iCLIP and PAR-CLIP [40, 41]. The application of such techniques to chromatin modifying complexes should be able to overcome some of the technical limitations in the analysis of lncRNA-protein associations.

RNA is a key to transcriptional activation

The first description of a lncRNA working as a co-activator molecule came as a surprise [10] and was later upheld by several reports [42-44]. Initial studies on noncoding RNAs with activating properties, like ncRNA-a7 and HOTTIP [10, 42], stemmed from the observation that siRNA-mediated depletion of the lncRNA resulted in loss of transcription at neighboring genes in a *cis* dependent manner. The specificity of the effect varies from regulation of a single gene by a lncRNA [10] to several genes residing in the same cluster [42]. Importantly, the distance between the lncRNA and the target gene can be as great as few hundreds kilobases and may encompass many additional loci that are not regulated by the lncRNA. Interestingly, a number of genes important for cellular proliferation are regulated by lncRNAs. It was shown that the lncRNAs termed ncRNA-a6 and ncRNA-a7 specifically regulate the expression of SNAIL2 and SNAIL1 genes, respectively [10]. Depletion of these lncRNAs phenocopied the knock-down of their target genes. For example, the knockdown of either ncRNA-a7 or SNAIL1 elicited similar defects in cellular migration [10]. Consistent with this observation, depletion of either the lncRNA or the target mRNA resulted in a similar gene expression signature as measured by microarray analysis [10]. The functional activity of lncRNAs could be experimentally dissected using reporter assays, similar to that used for interrogation of transcriptional enhancers. Using these assays one can demonstrate that it is the lncRNAs and not the act of transcription through the lncRNA locus per se that is required for the enhancer-like activity [10]. Such an assay has been a useful tool for dissection of the mechanism of action of enhancer RNAs [45-47].

A similar enhancer activity was shown for HOTTIP, a lncRNA adjoining the HOXA cluster. HOTTIP stimulated transcription of the nearest HOXA genes and the effect gradually declined with the distance from the HOTTIP locus. HOTTIP established an active chromatin state for nearly half of the HOXA cluster, facilitating recruitment of WDR5 and therefore orchestrating MLL-dependent methylation of H3K4 at the transcription start site of six HOXA genes [42]. Notably, chromosome conformation maps revealed an intricate network of interactions between the HOTTIP locus and the entire 50kb of adjacent chromatin that accommodated all six regulated HOXA genes [42]. While HOTTIP was not deemed a requirement for the DNA looping, other studies have suggested that coordinating physical interactions with the lncRNAs loci and their targets may be a common function of many lncRNAs [45, 47, 48]. It was shown that enhancer RNA, ncRNA-a3, regulated its target gene, TAL1, from a distance of about 50 kb [10]. Both ncRNA-a7 and ncRNA-a3 loci are found in long-range chromatin loops with their targets that are established by the Mediator complex [45]. It was reported that activating lncRNAs associated with the Mediator complex and through such interactions potentiated its kinase activity towards histone H3.1 [45]. LncRNA depletion by RNAi decreased Mediator's occupancy at their target genes, suggesting that lncRNAs tethered the Mediator complex to their target sites [45, 48, 49].

A number of other lncRNAs have recently been shown to display activating properties. Two outstanding examples are lincRNA-p21 and NeST [44, 50]. As previously discussed, lincRNA-p21 was originally described as a trans-acting RNA, adjacent to the p21 locus, capable of repressing p53 targets via hnRNP-K binding [36]. However, a recent report using mouse knock-outs of lincRNA-p21 has unveiled a moderate but dependable upregulation of p21 by its noncoding neighbor [50], suggesting that this direct regulation of the p21 gene precedes and controls the repressive effects on the p53 gene [36]. The lncRNA NeST is adjacent to a locus encoding for the IFN- γ gene in mice, which is also well conserved in the human genome where a similar genomic architecture can be observed. NeST RNA greatly enhanced IFN- γ transcription, specifically in CD8+ T cells upon infection [44]. Interestingly, while NeST is transcribed from a locus located approximately 50 kb downstream of the IFN- γ gene, its effect could be mimicked through over-expression in *trans* [44]. Therefore, while most enhancer RNAs seem to mediate their effects in a cis-dependent fashion, it is likely that in some cases their over-expression may provide a trans-activating function.

Another activating lncRNA, named TINCR, was recently identified as one of the most expressed lncRNAs during the transition from progenitors to differentiated keratinocytes [51]. The authors suggested that TINCR expression controlled an entire set of genes required for epidermal differentiation. However TINCR did not seem to mediate its effects through transcriptional regulation. Instead, TINCR interacted with the mature mRNAs of epidermal differentiation genes and improved their stability by means of an RNA-binding protein named STAU1 [51]. Intriguingly, a post-transcriptional function was also reported for lincRNA-p21 [52], although having a repressive function. RNAs like TINCR may represent yet another functional class of lncRNAs, whose action takes place predominantly in the cytoplasm and regulates stability and/or ribosomal access to large pools of messenger RNAs.

Enhancers in the spotlight

The discovery of a class of long noncoding RNAs with enhancer-like features not only spurred several questions about the biology of lncRNAs but also challenged our knowledge of enhancers as static DNA elements. Molecules such as ncRNA-a7, HOTTIP and lincRNA-p21 are capable of stimulating their target genes in a tissue and cell specific manner. These observations raised the critical question of whether lncRNAs define distal regulatory elements known as enhancers and whether enhancers are transcribed as lncRNAs.

Initial genome-wide studies of enhancers established a chromatin-based definition of enhancer regions [53-56]. A typical enhancer was defined to be enriched in mono-methylated H3K4 and acetylated H3K27 while diminished in tri-methylated H3K4. Enhancers were also shown to bind p300/CBP co-activators and constitute a binding platform for several sequence specific transcription factors (TF). Indeed, enhancers were traditionally defined as DNase I hypersensitive sites that are enriched in transcription factor binding sites, a feature they share with the proximal regulatory region of promoters of coding genes. Interestingly, chromatin signatures reported for most polyadenylated lncRNAs are closer to that of protein coding genes (rich in H3K4 trimethylation and low in monomethylation) than regions defined as distal regulatory sites (low in H3K4 trimethylation and high in monomethylation). Such distal regulatory sites were shown to express bidirectional transcripts that are predominantly non polyadenylated with low levels of splicing. It is not clear whether such differences in the ratio of high and low H3K4 mono- vs tri-methylation endow specific functional characteristics on a DNA regulatory element since most promoters were also shown to express bidirectional transcripts.

In the pre-genomic era, a handful of well-studied enhancers regions were known to be transcriptionally active. For instance, there was evidence of RNAPII binding and transcription occurring at the major histocompatibility complex [57]. Transcription was also observed at the locus control region (LCR) and within the intergenic regions of the beta-globin locus [58]. Moreover, DNA elements upstream of genes encoding for the T cell receptor [59] and the lysozyme [60] were also shown to be transcribed. It was not until 2010 that initial genome-wide analyses described widespread association of RNAPII in regions upstream of protein coding genes with enhancer-like signatures. RNAPII was dynamically recruited to these enhancers upon endotoxin stimulation of primary macrophages [61] or calcium-dependent activation of primary neuronal cultures [20]. Engagement of the RNAPII machinery at mammalian enhancers was later corroborated by genome-wide analysis of a complete set of general transcription factors (GTFs) [17]. Both of the initial reports described multiple well-defined sites of transcription and RNAPII recruitment extending for up to 50-60 kilobases upstream of genes like Fos, Arc and Ccl5 that are reminiscent of locus control regions and were later described as super-enhancer regions [62]. In human macrophages, the authors examined selected enhancer loci and reported uni-directional transcription (antisense with respect to the adjacent coding gene) of a few kb long RNA, which was unspliced and polyadenylated [61]. In the case of neuronal cells, the analysis was focused on a group of about 2,000 stimulus-dependent CBP binding sites that also display RNAPII binding [20]. RNA-seq revealed bi-directional transcription of non polyadenylated transcripts, originating from the midst of the CBP binding site. Induction of these enhancer-

derived RNAs (eRNAs) was correlated with induction of the neighboring genes, similar to that of activating lncRNAs [10, 42, 50].

Tackling the function of eRNAs

Enhancer RNAs have emerged as a common feature of active enhancer regions. The initial studies in macrophages and neurons were corroborated by additional reports of actively transcribed enhancers responding to FoxA1 in prostatic cells [63], to estrogen in breast cancer cells [47, 64] and to PPAR- γ in adipocytes [65]. These experiments utilized a newly developed technique of Global Run-On sequencing (GRO-seq), a genome-wide analysis which expanded the traditional run-on assays to glance at nascent RNA molecules as they are being extended by RNA polymerase II [66]. GRO-seq has been a tremendous asset in exploring eRNA biogenesis. Indeed, the transient nature of these transcripts is a key feature that delayed the discovery of a widespread transcription at enhancers. GRO-seq provided bioinformaticians with more dependable data sets allowing de novo identification of eRNAs [46, 47, 64] and established a powerful methodology to investigate whether their regulation was occurring at the initiation or elongation step of RNA polymerase II-directed transcription. In addition, GRO-seq protocols modified to select for methyl-guanosine capped transcripts allow for an even more accurate mapping of eRNA transcription start sites [46].

One of the major findings validated by GRO-seq data concerns the bi-directionality of enhancer RNAs. The DNase I hypersensitive regions at the core of the enhancers recruit RNAPII that initiates transcription in both directions on opposite strands [46, 47, 64]. The exact significance and implications of this bi-directional RNA synthesis at enhancers is currently unclear. It is interesting that a number of reports suggested that only one of the two strands may be functional in transcriptional activation [46-48]. However there are no discerning features that allow for the prediction of the functional strand of the eRNA. A further advancement in the mapping and annotation of human eRNAs came from the completion of the FANTOM5 project. Several hundreds of cells and tissues were analyzed for their chromatin landscape and transcriptome, resulting in a broader and deeper mapping of enhancer elements across the human genome. It was proposed that most enhancer RNAs are unspliced and shorter than previously reported (median length of 346bp) [67]. They were shown to be predominantly nuclear, non-polyadenylated, and methyl-guanosine capped, as indicated by CAGE analysis, which corroborated previous evidence from 5'-GRO-seq data [46]. A number of reports have shown that eRNAs were induced by binding of tissue-specific transcription factors [63, 68, 69], and upon induction they cooperated with the transcriptional co-activator complex Mediator [48, 65] and Cohesin [47] to stimulate transcription of their target genes. Similar to activating lncRNAs [45] the role of Mediator in eRNA function entailed the formation of long distance chromatin loops [48, 63] with the Cohesin complex providing additional scaffolding [47].

Enhancers are dynamic loci and their activation promotes the deposition of chromatin marks. It was shown that transcription of eRNAs promotes mono and di-methylation of lysine 4 residues on histone H3 [69], which is likely operated by Mll3 containing complexes [69, 70]. Conversely, repressed enhancers can lose their chromatin signature relatively fast

and acquire a status of 'latency' [68]. Latent enhancers can be promptly awakened in a transcription factor dependent manner, triggering nucleosome remodeling and regaining full RNAPII accessibility [68]. As was indicated earlier, induction of eRNAs has been associated with the activation of a neighboring genes in different model systems [20, 47, 61, 63, 64]. It was shown that a set of p53-regulated enhancers were transcribed in primary fibroblasts [71]. Importantly, the transcription of eRNAs was dependent on p53 binding. Using a luciferase reporter assay as the transcriptional read out, the authors tethered the eRNA to a reporter promoter and observe increased luciferase activity [71]. Additionally, implementation of 4C approach allowed the authors to demonstrate long-range interactions of an enhancer region with multiple distal loci. Importantly, siRNA depletion of the p53-induced enhancer RNAs significantly impaired activation of p53 target genes while there were no detected changes in chromosome looping.

In a related work on the master regulator of myogenic differentiation, MyoD, the authors described a regulatory feedback modulated by a set of eRNAs [72]. MyoD was shown to bind several thousand extragenic regions containing enhancer signatures. These regions exhibited divergent non-polyadenylated transcription typical of eRNAs, which was dependent on the binding of MyoD. Among the MyoD regulated enhancers was a large genomic region upstream of MyoD itself, containing a number of distinct eRNAs. Depletion of some of these eRNAs significantly reduced MyoD mRNA levels and significantly decreased RNAPII recruitment to the MyoD promoter [72]. The nuclear receptor Rev-Erb was the first example of an enhancer-binding factor that repressed the expression of eRNAs [46]. Rev-Erb was shown to turn off macrophage-specific enhancers by reducing their transcriptional activity. Transcription of macrophage enhancers, in turn, is the driving force behind the expression of inflammatory genes like *MMP9* and *CX3CR1*. Such transcriptional activation was dependent on the expression of eRNAs, since their depletion by antisense DNA oligos impaired inflammatory response. When subcloned into a gene-reporter plasmid, inflammatory eRNAs displayed typical enhancer features independent of their orientation. Transactivation required the full-length transcription of the noncoding RNA, while the mere act of initiation and early elongation did not translate into a significant enhancer response [46].

Another insightful example of eRNA-driven gene activation was described by the action of estrogen receptor (ER). A large fraction of ER binding sites in breast cancer cells were deemed extragenic and corresponded to bi-directionally transcribed eRNAs [47, 64]. Similar to the above-mentioned enhancers, depletion of the noncoding transcripts affected ER-mediated activation of the neighboring genes. ChIA-PET and 3D-DSL analyses revealed intricate chromatin looping occurring between eRNA sites and protein coding genes they regulate [47, 64]. It was shown that such three-dimensional architecture was coordinated by eRNAs, via recruitment of components of the Cohesin complex [47]. Taken together, these studies point to a role for eRNAs as dynamic molecular switches that can be activated and/or silenced within a short time span. DNA-binding motifs direct transcription factors to enhancer sites, promoting bidirectional transcription of eRNAs (Fig. 1). Chromatin-associated eRNAs promote DNA looping of the enhancer elements with a protein coding

locus residing in *cis* to control the tissue- and temporal-specific expression of the messenger RNA (Fig.1).

Perspectives

The prominent role of lncRNAs in tissue-specific expression of genes critical for cellular growth and differentiation has unveiled new and unanticipated layers of transcriptional regulation. The most intriguing aspect of lncRNAs and eRNAs is their prominent tissue and cell specificity, which largely exceeds that of protein coding genes. Moreover, tissue specificity of lncRNAs is not incidental but is maintained throughout evolution [73]. It is likely that lncRNAs play a prominent role during tissue development and organogenesis, as it has been suggested by studies in zebrafish [18] and mouse [74].

Although the discovery of eRNAs is reshaping our view on transcriptional regulation and enhancers, we are still in need of further insight into their mechanism of action. Despite a rapid progress following their initial identification, eRNAs remain poorly characterized, owing to their unstable nature and short half-life. It will be important to arrive at an improved annotation of eRNAs during different stages of developmental control and in disease states. Moreover, we are still at early stages regarding the pathways required for the biogenesis of primary eRNA transcripts to their mature state. Beside the clear-cut involvement of the RNAPII machinery and the absence of participation of splicing machinery in their synthesis, their maturation has not been explored. While some eRNAs are reported to be polyadenylated, the majority appear to be processed in a different manner than the mRNA genes.

Finally, the disease relevance of noncoding RNAs is going to generate a great deal of excitement in the near future. The mapping of thousands of lncRNAs and eRNAs leads to an important consequence: we now have the possibility to analyze genetic variations with a brand new functional point of view. There is already evidence of disease-causing SNPs associated with regulatory regions [75, 76]. Several comprehensive catalogs of active eRNAs have already been compiled and the map of risk-associated genetic variants keeps growing [77]. A systematic cross-comparison between these sets of data will be extremely valuable for both human genetics and noncoding RNA biology. Genetic diseases have helped to elucidate the functional domains of disease-causing proteins, given that mutations in the DNA sequence of a gene have often exposed critical domains or catalytic sites in poorly characterized polypeptides. Similarly, disease-inducing SNPs lying in regulatory regions may uncover functional domains and critical structures used by eRNAs to target the promoter of coding genes and regulate their transcription. With the advent of new technological developments in sequencing and mapping the three dimensional structure of the genome, we are poised to tackle the fundamental questions regarding how genomes are organized and how lncRNAs participate and organize genome architecture and its transcriptional output.

Acknowledgments

R.S. is supported by grant R01 GM078455 and R01 GM105754 from the National Institute of Health (NIH).

References

1. Birney, E. Stamatoyannopoulos; Dutta, JA.; Guigo, A.; Gingeras, R.; Margulies, TR.; Weng, EH.; Snyder, Z.; Dermitzakis, M.; Thurman, ET.; Kuehn, RE.; Taylor, MS.; Neph, CM.; Koch, S.; Asthana, CM.; Malhotra, S.; Adzhubei, A.; Greenbaum, I.; Andrews, JA.; Flicek, RM.; Boyle, P.; Cao, PJ.; Carter, H.; Clelland, NP.; Davis, GK.; Day, S.; Dhami, N.; Dillon, P.; Dorschner, SC.; Fiegler, MO.; Giresi, H.; Goldy, PG.; Hawrylycz, J.; Haydock, M.; Humbert, A.; James, R.; Johnson, KD.; Johnson, BE.; Frum, EM.; Rosenzweig, TT.; Karnani, ER.; Lee, N.; Lefebvre, K.; Navas, GC.; Neri, PA.; Parker, F.; Sabo, SC.; Sandstrom, PJ.; Shafer, R.; Vetrie, A.; Weaver, D.; Wilcox, M.; Yu, S.; Collins, M.; Dekker, FS.; Lieb, J.; Tullius, JD.; Crawford, TD.; Sunyaev, GE.; Noble, S.; Dunham, WS.; Denoeud, I.; Reymond, F.; Kapranov, A.; Rozowsky, P.; Zheng, J.; Castelo, D.; Frankish, R.; Harrow, A.; Ghosh, J.; Sandelin, S.; Hofacker, A.; Baertsch, IL.; Keefe, R.; Dike, D.; Cheng, S.; Hirsch, J.; Sekinger, HA.; Lagarde, EA.; Abril, J.; Shahab, JF.; Flamm, A.; Fried, C.; Hackermuller, C.; Hertel, J.; Lindemeyer, J.; Missal, M.; Tanzer, K.; Washietl, A.; Korbel, S.; Emanuelsson, J.; Pedersen, O.; Holroyd, JS.; Taylor, N.; Swarbreck, R.; Matthews, D.; Dickson, N.; Thomas, MC.; Weirauch, DJ.; Gilbert, MT.; J., et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007; 447:799–816. [PubMed: 17571346]
2. Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, Sementchenko V, Piccolboni A, Bekiranov S, Bailey DK, Ganesh M, Ghosh S, Bell I, Gerhard DS, Gingeras TR. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* (New York, NY. 2005; 308:1149–54.
3. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* (New York, NY. 2007; 316:1484–8.
4. Consortium, E. P. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]
5. Mouse EC, Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M, Bender M, Kaul R, Canfield T, Giste E, Johnson A, Zhang M, Balasundaram G, Byron R, Roach V, Sabo PJ, Sandstrom R, Stehling AS, Thurman RE, Weissman SM, Cayting P, Hariharan M, Lian J, Cheng Y, Landt SG, Ma Z, Wold BJ, Dekker J, Crawford GE, Keller CA, Wu W, Morrissey C, Kumar SA, Mishra T, Jain D, Byrsk-Bishop M, Blankenberg D, Lajoie BR, Jain G, Sanyal A, Chen KB, Denas O, Taylor J, Blobel GA, Weiss MJ, Pimkin M, Deng W, Marinov GK, Williams BA, Fisher-Aylor KI, Desalvo G, Kiralusha A, Trout D, Amrhein H, Mortazavi A, Edsall L, McCleary D, Kuan S, Shen Y, Yue F, Ye Z, Davis CA, Zaleski C, Jha S, Xue C, Dobin A, Lin W, Fastuca M, Wang H, Guigo R, Djebali S, Lagarde J, Ryba T, Sasaki T, Malladi VS, Cline MS, Kirkup VM, Learned K, Rosenbloom KR, Kent WJ, Feingold EA, Good PJ, Pazin M, Lowdon RF, Adams LB. An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome biology*. 2012; 13:418. [PubMed: 22889292]
6. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigo R, Gingeras TR. Landscape of transcription in human cells. *Nature*. 2012; 489:101–8. [PubMed: 22955620]
7. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhhattar R, Gingeras TR, Hubbard TJ,

- Notredame C, Harrow J, Guigo R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research*. 2012; 22:1775–89. [PubMed: 22955988]
8. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009; 458:223–7. [PubMed: 19182780]
 9. Lv J, Liu H, Huang Z, Su J, He H, Xiu Y, Zhang Y, Wu Q. Long non-coding RNA identification over mouse brain development by integrative modeling of chromatin and genomic features. *Nucleic Acids Res*. 2013; 41:10044–61. [PubMed: 24038472]
 10. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhhattar R. Long noncoding RNAs with enhancer-like function in human cells. *Cell*. 2010; 143:46–58. [PubMed: 20887892]
 11. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*. 2011; 147:789–802. [PubMed: 22056041]
 12. Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell*. 2013; 154:240–51. [PubMed: 23810193]
 13. Galindo MI, Pueyo JI, Fouix S, Bishop SA, Couso JP. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS biology*. 2007; 5:e106. [PubMed: 17439302]
 14. Kondo T, Plaza S, Zanet J, Benrabah E, Valenti P, Hashimoto Y, Kobayashi S, Payre F, Kageyama Y. Small peptides switch the transcriptional activity of Shavenbaby during *Drosophila* embryogenesis. *Science*. 2010; 329:336–9. [PubMed: 20647469]
 15. Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN, Ma J, Levin JZ, Karger AD, Budnik BA, Rinn JL, Saghatelian A. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nature chemical biology*. 2013; 9:59–64. [PubMed: 23160002]
 16. Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeier S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F, Kuster B. Mass-spectrometry-based draft of the human proteome. *Nature*. 2014; 509:582–7. [PubMed: 24870543]
 17. Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C, Eick D, Gut I, Ferrier P, Andrau JC. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nature structural & molecular biology*. 2011; 18:956–63.
 18. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. 2011; 147:1537–50. [PubMed: 22196729]
 19. Tilgner H, Knowles DG, Johnson R, Davis CA, Chakraborty S, Djebali S, Curado J, Snyder M, Gingeras TR, Guigo R. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome research*. 2012; 22:1616–25. [PubMed: 22955974]
 20. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010; 465:182–7. [PubMed: 20393465]
 21. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & development*. 2011; 25:1915–27. [PubMed: 21890647]
 22. Tan MH, Au KF, Yablonovitch AL, Wills AE, Chuang J, Baker JC, Wong WH, Li JB. RNA sequencing reveals a diverse and dynamic repertoire of the *Xenopus tropicalis* transcriptome over development. *Genome research*. 2013; 23:201–16. [PubMed: 22960373]

23. Aanes H, Winata CL, Lin CH, Chen JP, Srinivasan KG, Lee SG, Lim AY, Hajan HS, Collas P, Bourque G, Gong Z, Korzh V, Alestrom P, Mathavan S. Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. *Genome research*. 2011; 21:1328–38. [PubMed: 2155364]
24. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ, Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE. The developmental transcriptome of *Drosophila melanogaster*. *Nature*. 2011; 471:473–9. [PubMed: 21179090]
25. Nam JW, Bartel DP. Long noncoding RNAs in *C. elegans*. *Genome research*. 2012; 22:2529–40. [PubMed: 22707570]
26. Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes & development*. 2012; 26:2392–407. [PubMed: 23073843]
27. Gendrel AV, Heard E. Noncoding RNAs and Epigenetic Mechanisms During XChromosome Inactivation. *Annu Rev Cell Dev Biol*. 2014
28. Barlow DP. Genomic imprinting: a mammalian epigenetic discovery model. *Annu Rev Genet*. 2011; 45:379–403. [PubMed: 21942369]
29. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007; 129:1311–23. [PubMed: 17604720]
30. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. Long noncoding RNA as modular scaffold of histone modification complexes. *Science*. 2010; 329:689–93. [PubMed: 20616235]
31. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, Abo R, Tabebordbar M, Lee RT, Burge CB, Boyer LA. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*. 2013; 152:570–83. [PubMed: 23352431]
32. da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, Sanulli S, Chow J, Schulz E, Picard C, Kaneko S, Helin K, Reinberg D, Stewart AF, Wutz A, Margueron R, Heard E. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Molecular cell*. 2014; 53:301–16. [PubMed: 24462204]
33. Kaneko S, Bonasio R, Saldana-Meyer R, Yoshida T, Son J, Nishino K, Umezawa A, Reinberg D. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Molecular cell*. 2014; 53:290–300. [PubMed: 24374312]
34. Jeon Y, Lee JT. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*. 2011; 146:119–33. [PubMed: 21729784]
35. 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 & development*. 2010; 24:2543–55. [PubMed: 20966046]
36. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M, Attardi LD, Regev A, Lander ES, Jacks T, Rinn JL. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*. 2010; 142:409–19. [PubMed: 20673990]
37. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nature structural & molecular biology*. 2013; 20:1250–7.
38. Cifuentes-Rojas C, Hernandez AJ, Sarma K, Lee JT. Regulatory interactions between RNA and polycomb repressive complex 2. *Molecular cell*. 2014; 55:171–85. [PubMed: 24882207]
39. Kaneko S, Son J, Bonasio R, Shen SS, Reinberg D. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes & development*. 2014; 28:1983–8. [PubMed: 25170018]

40. Sugimoto Y, Konig J, Hussain S, Zupan B, Curk T, Frye M, Ule J. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome biology*. 2012; 13:R67. [PubMed: 22863408]
41. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*. 2010; 141:129–41. [PubMed: 20371350]
42. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*. 2011; 472:120–4. [PubMed: 21423168]
43. Bertani S, Sauer S, Bolotin E, Sauer F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Molecular cell*. 2011; 43:1040–6. [PubMed: 21925392]
44. Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, Chang HY, Brahic M, Kirkegaard K. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. *Cell*. 2013; 152:743–54. [PubMed: 23415224]
45. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature*. 2013; 494:497–501. [PubMed: 23417068]
46. Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, Lee CY, Watt A, Grossman TR, Rosenfeld MG, Evans RM, Glass CK. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature*. 2013; 498:511–5. [PubMed: 23728303]
47. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature*. 2013; 498:516–20. [PubMed: 23728302]
48. Hsieh CL, Fei T, Chen Y, Li T, Gao Y, Wang X, Sun T, Sweeney CJ, Lee GS, Chen S, Balk SP, Liu XS, Brown M, Kantoff PW. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111:7319–24. [PubMed: 24778216]
49. Trimarchi T, Bilal E, Ntziachristos P, Fabbri G, Dalla-Favera R, Tsiganos A, Aifantis I. Genome-wide Mapping and Characterization of Notch-Regulated Long Noncoding RNAs in Acute Leukemia. *Cell*. 2014; 158:593–606. [PubMed: 25083870]
50. Dimitrova N, Zamudio JR, Jong RM, Soukup D, Resnick R, Sarma K, Ward AJ, Raj A, Lee JT, Sharp PA, Jacks T. LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint. *Molecular cell*. 2014; 54:777–90. [PubMed: 24857549]
51. Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J, Johnston D, Kim GE, Spitale RC, Flynn RA, Zheng GX, Aiyer S, Raj A, Rinn JL, Chang HY, Khavari PA. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*. 2013; 493:231–5. [PubMed: 23201690]
52. 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. *Molecular cell*. 2012; 47:648–55. [PubMed: 22841487]
53. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics*. 2007; 39:311–8. [PubMed: 17277777]
54. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011; 470:279–83. [PubMed: 21160473]
55. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory

- elements required for macrophage and B cell identities. *Molecular cell*. 2010; 38:576–89. [PubMed: 20513432]
56. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenko VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009; 459:108–12. [PubMed: 19295514]
 57. Masternak K, Peyraud N, Krawczyk M, Barras E, Reith W. Chromatin remodeling and extragenic transcription at the MHC class II locus control region. *Nat Immunol*. 2003; 4:132–7. [PubMed: 12524537]
 58. Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ. Intergenic transcription and transinduction of the human beta-globin locus. *Genes & development*. 1997; 11:2494–509. [PubMed: 9334315]
 59. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor alpha recombination. *EMBO J*. 2007; 26:4380–90. [PubMed: 17882258]
 60. Lefevre P, Witham J, Lacroix CE, Cockerill PN, Bonifer C. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Molecular cell*. 2008; 32:129–39. [PubMed: 18851839]
 61. De Santa F, Barozzi I, Miettinen F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS biology*. 2010; 8:e1000384. [PubMed: 20485488]
 62. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. *Cell*. 2013; 155:934–47. [PubMed: 24119843]
 63. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, Glass CK, Rosenfeld MG, Fu XD. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature*. 2011; 474:390–4. [PubMed: 21572438]
 64. Hah N, Murakami S, Nagari A, Danko CG, Kraus WL. Enhancer transcripts mark active estrogen receptor binding sites. *Genome research*. 2013; 23:1210–23. [PubMed: 23636943]
 65. Step SE, Lim HW, Marinis JM, Prokesch A, Steger DJ, You SH, Won KJ, Lazar MA. Anti-diabetic rosiglitazone remodels the adipocyte transcriptome by redistributing transcription to PPARgamma-driven enhancers. *Genes & development*. 2014; 28:1018–28. [PubMed: 24788520]
 66. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*. 2008; 322:1845–8. [PubMed: 19056941]
 67. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jorgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhashi E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Muller F, Forrest AR, Carninci P, Rehli M, Sandelin A. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014; 507:455–61. [PubMed: 24670763]
 68. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, Curina A, Prosperini E, Ghisletti S, Natoli G. Latent enhancers activated by stimulation in differentiated cells. *Cell*. 2013; 152:157–71. [PubMed: 23332752]
 69. Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, Chun HB, Tough DF, Prinjha RK, Benner C, Glass CK. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Molecular cell*. 2013; 51:310–25. [PubMed: 23932714]
 70. Herz HM, Mohan M, Garruss AS, Liang K, Takahashi YH, Mickey K, Voets O, Verrijzer CP, Shilatifard A. Enhancer-associated H3K4 monomethylation by Trithorax-related, the *Drosophila* homolog of mammalian Mll3/Mll4. *Genes & development*. 2012; 26:2604–20. [PubMed: 23166019]

71. Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, Elkon R, Melo SA, Leveille N, Kalluri R, de Laat W, Agami R. eRNAs are required for p53-dependent enhancer activity and gene transcription. *Molecular cell*. 2013; 49:524–35. [PubMed: 23273978]
72. Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartorelli V. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Molecular cell*. 2013; 51:606–17. [PubMed: 23993744]
73. Washietl S, Kellis M, Garber M. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. *Genome research*. 2014; 24:616–28. [PubMed: 24429298]
74. Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, Liapis SC, Mallard W, Morse M, Swerdel MR, D'Ecclesiss MF, Moore JC, Lai V, Gong G, Yancopoulos GD, Frendewey D, Kellis M, Hart RP, Valenzuela DM, Arlotta P, Rinn JL. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife (Cambridge)*. 2013; 2:e01749.
75. Cowper-Salari R, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoutte J, Moore JH, Lupien M. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nature genetics*. 2012; 44:1191–8. [PubMed: 23001124]
76. Zhang X, Cowper-Salari R, Bailey SD, Moore JH, Lupien M. Integrative functional genomics identifies an enhancer looping to the SOX9 gene disrupted by the 17q24.3 prostate cancer risk locus. *Genome research*. 2012; 22:1437–46. [PubMed: 22665440]
77. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. *American journal of human genetics*. 2013; 93:779–97.1. [PubMed: 24210251]

Figure 1. Mechanism of transcriptional activation mediated by enhancer RNAs

The model illustrates different regulatory steps that lead to the activation of a protein coding gene. A typical p300/CBP-dependent enhancer is located distal from its target protein coding locus (I) and depends on p300/CBP to sustain H3K27 acetylation and maintain an open chromatin state. Binding of a sequence specific transcription factor (II) recruits the Mediator complex (Med) to the enhancer, which engages the RNAPII machinery (III) to initiate transcription on both strands. The short-lived eRNA transcripts remain associated with chromatin and bind to Mediator/Cohesin to help enforce a DNA looping between the enhancer and the distal gene (IV). Such chromatin conformation changes culminate in the transcriptional activation of the target gene, resulting in the production of mature messenger RNAs that can be exported to the cytoplasm (IV).